THE EFFICACY OF TREATMENTS FOR CHILDHOOD DEPRESSION:

AN INTEGRATIVE REVIEW

by

Kurt David Michael

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Psychology

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1999
Prevalence estimates of child depression are substantial and morbidity leads to an increased risk of recurrence during adolescence and adulthood. Further, early-onset depression is associated with a number of negative outcomes including: poor physical health, social and interpersonal impairments, academic problems, substance abuse, future maladjustment, and suicidal behavior. In light of the prevalence, persistence, and negative outcomes associated with depression in children and adolescents, several treatments ranging from psychosocial to pharmacological interventions have been developed and evaluated. However, the overall efficacy of treatments remains equivocal because the majority of existing reviews of the child and adolescent depression treatment literature are narrative in nature, methodologically flawed, and/or present vague or conflicting conclusions. Although there are a number of good meta-analytic reviews that indicate that psychotherapy is effective with children and adolescents overall,
comprehensive meta-analytic reviews focusing on the efficacy of psychological treatments specifically for depressed youth are nonexistent in the published literature.

A comprehensive sample of studies on the psychosocial and pharmacological treatment of early-onset depression was located through an extensive literature search. Articles that met the inclusionary criteria were subsequently analyzed. The outcome data from 37 outcome studies were extracted and converted into effect sizes. Comparisons of main effects, potential interactions, and other specified variables were conducted. The overall findings of this meta-analysis indicate that several different psychosocial interventions for early-onset depression produce moderate to large treatment gains that are clinically meaningful for many afflicted youth. Further, it appears that psychosocial treatments are, in general, superior to pharmacological regimens in treating depressed children and adolescents. However, there is also recent evidence that selective serotonin reuptake inhibitors such as fluoxetine are efficacious, and they will likely play an increased role in the management of affective illness in youngsters. The clinical implications and limitations of these data are discussed and suggestions for future research are provided.
DEDICATION

I dedicate this dissertation to my wife, Amy, and my son, Kauner. I offer my most sincere acknowledgment to my wife as she endured a great deal during the course of this project including my multiple responsibilities, my bizarre nocturnal behaviors, and my mood swings. Most importantly, my wife assumed the majority role in the care of our most precious blessing, our son. For this, and for all of the other wonderful things that she does, I will be forever grateful.
ACKNOWLEDGMENTS

I offer my most sincere appreciation to Dr. Susan Crowley for her support and guidance throughout this dissertation project. From the project's genesis through the final defense, Dr. Crowley provided me with encouragement, insight, and analytical advice to make this the best possible document. In addition, I thank Dr. David Stein, Dr. Pat Truhn, Dr. Tamara Ferguson, and Dr. Dennis Odell for serving on my committee and for their thoughtful feedback and assistance throughout the course of the study. Moreover, I thank Jim Sharpnack, a close friend and colleague, who double-coded each of my studies, which was essential for computing intercoder reliability estimates. I am also grateful for Dr. Karl White's technical suggestions regarding coding, calculation, and interpretive guidelines for meta-analytic designs. Finally, I extend my sincere gratitude to Karen Ranson for her assistance in the preparation of this manuscript.

Kurt David Michael
# CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>DEDICATION</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGMENTS</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>I</td>
<td>INTRODUCTION AND PROBLEM STATEMENT</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>REVIEW OF THE LITERATURE</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Depression in Children</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Treatment of Childhood Depression</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Treatment Variables</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Purpose and Objectives</td>
<td>57</td>
</tr>
<tr>
<td>III</td>
<td>METHODOLOGY</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Population and Sample</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Design</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Instrumentation and Analysis</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>70</td>
</tr>
<tr>
<td>IV</td>
<td>RESULTS</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Descriptive Statistics</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Intercoder Agreement</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Main Effects</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Analysis of Study Variables</td>
<td>81</td>
</tr>
<tr>
<td>V</td>
<td>DISCUSSION</td>
<td>98</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Main Effects</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>Study Variables</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Summary and Clinical Implications</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>Limitations and Future Research</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>APPENDIX</td>
<td>144</td>
<td></td>
</tr>
<tr>
<td>VITA</td>
<td>147</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Frequency and Percentage of Study Characteristics</td>
<td>73</td>
</tr>
<tr>
<td>2. Intercoder Agreement and Cohen's Kappa of Study Variables</td>
<td>76</td>
</tr>
<tr>
<td>3. Overall Mean Effect Sizes at Posttreatment</td>
<td>77</td>
</tr>
<tr>
<td>4. Overall Mean Effect Sizes at Follow-Up</td>
<td>77</td>
</tr>
<tr>
<td>5. Effect sizes at Posttreatment and Follow-Up for Controlled Psychosocial Studies</td>
<td>78</td>
</tr>
<tr>
<td>6. Effect Sizes at Posttreatment and Follow-Up for Each Pre- and Postpsychosocial Study</td>
<td>79</td>
</tr>
<tr>
<td>7. Effect Sizes Comparing Active Versus Placebo Positive Rates for Controlled Pharmacological Studies</td>
<td>82</td>
</tr>
<tr>
<td>8. Mean Effect Sizes Based on Average Age of Sample</td>
<td>84</td>
</tr>
<tr>
<td>9. Interaction Between Effect Sizes and Sex in Psychosocial Studies</td>
<td>86</td>
</tr>
<tr>
<td>10. Interaction Between Effect Sizes and Referral Source in Psychosocial Studies</td>
<td>87</td>
</tr>
<tr>
<td>11. Mean Effect Sizes Based on Treatment Setting</td>
<td>88</td>
</tr>
<tr>
<td>12. Mean Effect Sizes by Treatment Modality in Controlled Psychosocial Studies</td>
<td>88</td>
</tr>
<tr>
<td>13. Mean Effect Sizes by Treatment Modality in Pre- and Poststudies</td>
<td>89</td>
</tr>
<tr>
<td>14. Mean Effect Sizes by Medication in Pharmacological Studies</td>
<td>90</td>
</tr>
<tr>
<td>15. Within-Study Effect Sizes Comparison of Assessment Measures</td>
<td>92</td>
</tr>
</tbody>
</table>
16. Mean Effect Sizes by Level of Training in Psychosocial Studies ........................................ 94

17. Mean Effect Sizes by Level of Severity in Psychosocial Studies ........................................ 96
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. DSM-IV (APA, 1994) symptom overlap: major depressive episode and generalized anxiety disorder</td>
<td>16</td>
</tr>
<tr>
<td>2. Relationship between study quality and standard mean difference effect sizes</td>
<td>97</td>
</tr>
</tbody>
</table>
Although childhood depression was not a recognized phenomenon until recently, it is now considered to be an important area of research in child psychopathology. Prior to its recognition, various researchers and clinicians questioned whether depression as a disorder could develop in children (Rie, 1966), or as described depressive equivalents (e.g., “masked depression”) that were purportedly manifested by overt behavioral problems such as delinquency, hyperactivity, and aggression (Cytryn & McKnew, 1972; Glaser, 1967). However, based on current research, depressive disorders in children are indeed real phenomena (Kovacs, 1997). Depression in children, for the most part, is characterized and identified in many of the same ways as depression in adults (American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders [4th ed.], 1994; Kovacs, 1996), and it adversely affects the lives of many youth.

Estimates of the prevalence of child and adolescent depression in the general population are substantial, ranging from 0.4% to 8.3% (Anderson, Williams, McGee, & Silva, 1987; Birmaher et al., 1996b; Fleming & Offord, 1990; Kashani et al., 1983; Lewinsohn, Clarke, Seeley, & Rohde, 1994). Among groups of clinically referred children, the prevalence rates are even higher (Poznanski & Mokros, 1994). Early-onset depression is often persistent and leads to an increased risk of recurrence during adolescence and adulthood (Kovacs, Obrosky, Gatsonis, & Richards, 1997). Further, depression in children and adolescents is associated with negative outcomes including: poor physical health, social and interpersonal impairments, academic problems, substance
abuse, future maladjustment, and suicidal behavior.

In light of the prevalence, persistence, and negative outcomes associated with depressive disorders in children, several researchers and clinicians have developed and implemented various treatments for childhood depression ranging from psychosocial interventions to pharmacological interventions.

Although there are a number of good meta-analytic reviews that indicate that psychotherapy is generally effective with children and adolescents who present with a broad array of problems (e.g., Casey & Berman, 1985; Weisz, Weiss, Han, & Granger, 1995), comprehensive reviews investigating the efficacy of psychosocial treatments for child and adolescent depression are currently nonexistent in the published literature.

In terms of pharmacological interventions for depressed youth, a number of open and controlled clinical trials have been conducted. In addition, the prescription of psychotropic medications for depression in children and adolescents has become a widely accepted practice. However, the efficacy and safety of such practices remain the subject of much debate and controversy, due to methodological problems, developmental considerations, and limited empirical support. Furthermore, there is a paucity of data on the comparison of pharmacological and psychosocial interventions for child and adolescent depression.

Current reviews of the outcome studies for childhood depression have not been adequate for guiding research, policy, and practice because many of the reviews are narrative in nature, methodologically flawed, and/or present vague or conflicting conclusions. Furthermore, while some of the efforts to review and synthesize the body of
literature on the treatments for depressed youth have answered some questions, there are many more questions left unanswered. From clinical and empirical perspectives, it is important to know whether outcomes covary with study and sample characteristics. For instance, which types of treatments are the best and for whom? Does the age of the child matter? Do comorbid conditions impact the outcomes? Does it make a difference whether the children have been identified as suffering from depressive symptomatology as opposed to a depressive disorder? Are the findings from a school-based sample generalizable to a clinic-based sample? What is the overall efficacy of pharmacological treatments for child and adolescent depression? How do psychosocial interventions compare with pharmacological treatments? Is there evidence of differential efficacy when a number of important variables are considered (e.g., age, sex, severity, methodological factors)?

If a comprehensive, methodologically sound review of the literature on the efficacy of psychosocial and pharmacological treatments for childhood depression existed, researchers and clinicians would have a better understanding of which treatments are effective in ameliorating depressive symptomatology in children as well as how outcomes covary with study and sample characteristics. Furthermore, a more comprehensive review would help to guide both intervention efforts and future empirical studies on the treatment of depressive disorders in children. Therefore, in the present study, an integrative review of early-onset depression treatments was conducted.
CHAPTER II
REVIEW OF RELATED LITERATURE

The present literature review was conducted to establish support for the need to complete a comprehensive, methodologically sound review of the outcome studies on the psychosocial and pharmacological treatments of childhood depression. A definition and brief history of childhood depression are presented, followed by a review of the prevalence, longitudinal course, correlates, and the assessment and diagnosis of childhood depressive disorders. For purposes of clarity, the term “children” is used inclusively subsuming children and adolescents between 6 and 18 years of age. However, when age-related distinctions are made, young children are referred to as those who are between 6 and 12 years old, whereas adolescents are identified as those who are between the ages of 13 and 18. Finally, a review of the current knowledge regarding the psychological and pharmacological treatment of childhood depression is included.

Depression in Children

As a symptom, depression is characterized by a dysphoric or unhappy mood state. Defined as a syndrome, depression consists of a constellation of behavioral and emotional symptoms that do not simultaneously exist by chance (Rehm & Tyndall, 1993). For example, when depressed mood is combined with labored psychomotor functioning, cognitive difficulties, and a lack of motivation, these symptoms, if experienced simultaneously, are often construed as evidence of a depressive syndrome. Establishing
the existence of a depressive disorder depends largely upon the duration and severity of
the symptoms as well as whether there is a pattern of marked impairment in important
areas of functioning (e.g., academic, social, physical).

From a historical perspective, there have been a limited number of empirical
efforts to investigate depressive illnesses in children because until recently, most
researchers and clinicians debated about the nature and existence of childhood depression
(Kaslow & Rehm, 1991). Although today there is a general consensus among researchers
and clinicians that childhood depression is similar to adult depression, three historical
conceptualizations of depression in children contributed to the lack of previous empirical
efforts in the area. First, based on the classic psychoanalytic viewpoint that was
predominant during the early part of this century, depression could not exist in children
because they lacked the mental structure (i.e., superego) necessary for the development of
depression. Rie (1966) pointed out that a number of psychoanalysts believed that
depression resulted from the discrepancy between the real self and the ideal self.
However, based on this conceptualization, a stable self-representation does not develop
until adolescence, and thus the major components of depression cannot develop during
childhood.

The second conceptualization of childhood depression, popular during the 1960s,
included depressive equivalents or “masked depression,” which were supposedly
manifested by a variety of behavioral problems such as temper tantrums, hyperactivity,
running away, learning disabilities, enuresis, and school failure (Cytryn & McKnew,
1972; Glaser, 1967). The major criticism of this conceptualization was the fact that
virtually any behavioral problem could be construed as masked depression, which
seriously hampered a clinician’s ability to discriminate between depressive and
nondepressive behavioral problems (Carlson & Cantwell, 1980). The third
conceptualization of childhood depression worthy of mention was the notion that
depressive symptoms (e.g., depressed mood) in children were merely transitory and not
indicative of a depressive disorder (Lefkowitz & Burton, 1978). This position was most
popular prior to the publication of American Psychiatric Association, Diagnostic and
Statistical Manual of Mental Disorders (3rd ed.) in 1980, in which formally diagnosing
children with depressive disorders was sanctioned by a number of professional groups
(DSM-III, 1980). However, based on several longitudinal studies of children with
depressive disorders, it is clear that these conditions are often more chronic and persistent
than previously thought (e.g., Fleming, Boyle, & Offord, 1993; Hughes, Preskorn, Wrona,
Hassanein, & Tucker, 1990; McCauley et al., 1993; McGee & Williams, 1988).

Despite the aforementioned conceptualizations that challenged the legitimacy of
depression in children, Kovacs (1989) suggested that there is now compelling evidence
from several empirical studies that children suffer from depression, regardless of whether
it is defined as a symptom (depressed mood), a syndrome (constellation of symptoms), or
a disorder (diagnosis based on the severity and duration of symptoms) as reflected in
current nosological systems such as the DSM-IV, (1994).

Prevalence and Longitudinal Course

Prevalence estimates of child and adolescent depression are varied, ranging from
0.4% to 8.3% (Birmaher et al., 1996b; Garrison et al., 1997; Lewinsohn et al., 1994; Silver, 1988). Further, the prevalence estimates for depression in child samples are reportedly lower than in adolescent cohorts (Harrington, Rutter, & Fombonne, 1996). However, it remains unclear whether the lower prevalence rates in child samples are due to actual differences in prevalence or an artifact of how developmental differences (e.g., language, self-understanding) might impact the assessment of depressive disorders in children (Cicchetti & Schneider-Rosen, 1986). In terms of sex differences in early-onset depressive disorders, it appears that males and females are equally affected during childhood (Birmaher et al., 1996b). However, as young children move into adolescence, the prevalence rates begin to shift whereby adolescent females experience depressive sequelae at twice the rate of their male counterparts, a sex ratio that closely resembles that of adult samples (Cicchetti & Toth, 1998).

In a critical evaluation of epidemiological studies of childhood depression, Fleming and Offord (1990) suggested that the variation in prevalence rates is due primarily to methodological flaws in the research designs including sampling bias, small samples, inconsistent measurements, age parameters, and diagnostic procedures. In light of these problems, the authors suggested that it is difficult to draw firm conclusions about the prevalence of childhood depression (Fleming & Offord, 1990). Nonetheless, even the lowest estimates are substantial, especially when one considers the long-term risk of depressive disorders in children and adolescents. For example, Kessler et al. (1994) reported that the lifetime prevalence of major depressive disorder (MDD) during adolescence is 15-20%, an estimate similar to adult samples.
The longitudinal course of childhood depression has received increased attention in the empirical literature over the last several years. Early studies characterized depressive disorders as transient, normal developmental reactions to stress and viewed them as unstable when compared to the relative stability of behavioral disorders (Fischer, Rolf, Hasazi, & Cummings, 1984; Graham & Rutter, 1973; McGee et al., 1985), especially conduct disorder (Offord et al., 1992). Clarizio (1984) posed the question of whether childhood depression is a chronic, transitory, or recurring condition. According to Clarizio, the answer depends on whether the sample under investigation is based upon a normal or clinical population, the severity and type of the depression, and how long the sample is followed.

In addition, determining the longitudinal course of depression depends upon whether one is measuring depressive symptomatology or attempting to determine whether a depressive disorder is present. The specific depressive symptomatology can vary over time and the determination of whether a disorder is present may reflect a diverse array of symptom clusters that change in frequency and severity above and below a particular diagnostic threshold. Kazdin (1990) described this phenomenon as the distinction between dimensional (symptomatology) versus categorical (diagnostic) assessment. Despite the aforementioned considerations, several recent longitudinal studies have provided evidence to support the idea that depressive disorders are more persistent than previously thought (Kovacs, 1989).

In one of the first longitudinal investigations of depressive disorders in children, Kovacs et al. (1984) reported that a significant number of youngsters remained
symptomatic for 5 years or more, even when treatment was implemented. In addition, McGee and Williams (1988) reported that 31% of depressed 9-year-olds were found to have persistent depressive disorders after 2- and 4-year follow-ups, at the ages of 11 and 13, respectively. In another longitudinal study, Nolen-Hoeksema, Girgus, and Seligman (1992) found that in a large sample of children who initially met criteria for serious levels of depression during initial testing, approximately 40% remained at that level for 6 months to 2 years.

Furthermore, DuBois, Felner, Bartels, and Silverman (1995) investigated the course and stability of self-reported depressive symptoms in a community sample of 435 school-age children. The authors reported that upon initial assessment, 10% of the sample presented with clinically significant symptoms of depression. Two years later, the authors reported that 32% of the "clinical" sample continued to endorse clinically significant levels of depressive symptomatology and evidenced a greater pattern of impairment across several areas of functioning. A number of other longitudinal investigations have augmented the notion that depressive disorders in children and adolescents are often characterized by a persistent, chronic course (e.g., Fleming et al., 1993; Kovacs, Akiskal, Gastonis, & Parrone, 1994; Kovacs et al., 1997; Lewinsohn et al., 1994). Thus, the findings from a number of empirical studies appear to contradict the belief that young children only suffer from brief and episodic depression, highlighting the importance of providing effective treatments early during the course of the illness.
Correlates of Depression in Children

Several researchers have asserted that depressive disorders in children typically result in significant distress, misery, and negative outcomes in young people including: diminished self-esteem (Kazdin, 1988), poor physical health (Costello et al., 1988), family dysfunction (Kashani, Burbach, & Rosenberg, 1988), increased risk for substance abuse (Kovacs, Goldston, & Gatsonis, 1993), future psychological problems (Kovacs, 1985), and a substantial risk for morbidity and mortality across the lifespan (Fleming & Offord, 1990; Harrington & Vostanis, 1995). There is compelling evidence that children and adolescents who suffer from depression have a variety of problems in daily living, ranging from unsatisfying interpersonal relationships (Puig-Antich et al., 1985) to noteworthy declines in academic performance (Fleming, Offord, & Boyle, 1989), and an increased risk for dropping out of school (Fleming & Offord, 1990).

Furthermore, Puig-Antich et al. (1985) reported that depressed children had more severe family, peer-related, and social problems when compared to nondepressed controls. Kovacs (1997) suggested that “the social problems of depressed youngsters may be related to or derive from their interpersonal behavior and impact on others” (p. 288). That is, impaired social functioning, which is often associated with depression, may both contribute to and be caused by depressive sequelae. For example, in an observational study of children in a school setting, Altmann and Gotlib (1988) reported that while depressed children attempted to engage in social exchanges, they spent considerable amounts of time alone, thus limiting the amount of potentially reinforcing social contact and the opportunities to practice and develop age-appropriate social skills (Kovacs &

In studies where the relationship between depression and impaired family functioning has been examined, greater amounts of conflict (i.e., marital, sibling, parent-child) have been reported in families with depressed versus nondepressed youngsters (e.g., Beck & Rosenberg, 1986; Kaslow, Rehm, Pollack, & Siegel, 1988). In addition, Kovacs (1997) suggested that the developmental, interpersonal, and social deficits associated with childhood depressive illnesses can have “ominous consequences in the context of family life by disrupting the attachment bond between parent and child” (p. 289).

Another reason to be concerned about child and adolescent depression is the association between clinical mood disturbances and suicide (Bettes & Walker, 1986; Rao, Weissman, Martin, & Hammond, 1993). Kovacs et al. (1993) reported that depressive disorders were associated with significantly higher rates of suicide than were nondepressive disorders in a mixed sample of children between the ages of 8 and 13. In addition, Kovacs et al. (1993) found that “in the presence of affective disorders, comorbid conduct and/or substance abuse disorders further increased the risk of suicide attempts” (p. 8).

In sum, depressive sequelae in children and adolescents are correlated with a plethora of impairments in developmental, intrapersonal, social, academic, and familial domains. These impairments, combined with the chronicity and likelihood of recurrence
of early-onset depression, provide ample justification for the need to develop, implement, and evaluate various interventions for depressed youth.

Assessment and Diagnosis

The efficacious treatment of virtually every psychological disorder is contingent upon an accurate assessment of the pathognomonic symptoms of that disorder (Achenbach, 1985). In addition, the process of assessment can have a profound impact on treatment outcome studies because the assessment and diagnostic procedures are used to (a) guide the inclusion/exclusion of subjects and (b) provide the basis for pretreatment, posttreatment, and follow-up evaluations of treatment effects (i.e., dependent measures). Thus, it is crucial to attend to these assessment issues when attempting to accurately assess, diagnose, and ultimately treat child psychological sequelae. In the case of childhood depression, the historical lack of consensus regarding its nature and existence has hindered the establishment of consistent diagnostic criteria and has impeded the development of treatment regimens. However, when the American Psychiatric Association (1980, 1987) revised the DSM-III, DSM-IIIR during the 1980s, the diagnosis of depressive disorders in children was formally sanctioned.

In the most recent version of the DSM-IV (APA, 1994), a diagnosis of either a depressive episode or a depressive disorder depends upon whether a minimum number of criteria (i.e., 5 out of 9) have been met. The criteria are made up of several emotional, cognitive, and behavioral symptoms, such as depressed mood, difficulty concentrating, changes in appetite, sleep disturbance, fatigue, thoughts of death, and impaired academic,
interpersonal, and social functioning. While the authors of the DSM-IV (APA, 1994) suggested that

the core symptoms of a Major Depressive Episode are the same for children and adolescents [as they are for adults]...certain symptoms such as somatic complaints, irritability, and social withdrawal are particularly common in children. (p. 324)

Furthermore, Carlson and Garber (1986) suggested that developmental factors (e.g., language, cognitive abilities, emotional insight) must be taken into account when attempting to identify and classify childhood depression. Nonetheless, Kovacs (1996) reported that there is evidence that depressive disorders “in childhood and in the later years of the life span capture the same psychopathological entity” (p. 711).

Despite more consistent and professionally sanctioned diagnostic criteria, the accurate assessment and diagnosis of depressive disorders in children is complicated by several factors. First, there are a number of different methods for assessing depressive symptomatology in children including direct behavioral observation, behavior rating scales, sociometric approaches, clinical interviews, and self-report measures. Each of these methods has limitations. For example, Merrell (1994) reported that the evaluation of childhood depression has typically relied upon the verbal or written reports of parents, teachers, and other significant figures in the child’s environment “due to the supposedly questionable accuracy of information obtained through self-report methods” (p. 194).

However, because depressive disorders are, in great measure, subjective perceptions of internal distress, they are often not readily or reliably identified by external observers (La Greca, 1990) and there is often disagreement among informants (Achenbach,
McConaughy, & Howell, 1987; Sanford, Offord, Boyle, Peace, & Racine, 1992). At the same time, the potential limitations of self-report measures include biased response styles (e.g., social desirability; Borg & Gall, 1989), children’s inability to understand and report their emotions (Clarizio, 1984), and whether the instrument has an age-appropriate reading level (Prout & Chizik, 1988). In light of these issues, La Greca (1990) emphasized the importance of a multimethod-multisource assessment that takes into account, “the limitations inherent in any one procedure” (p. 8). However, from a historical perspective, assessment protocols utilizing multiple methods and multiple sources across settings have been the exception rather than the rule (Merrell, 1994).

A second factor that complicates the assessment and diagnosis of depression in children is the fact that depressive disorders are a heterogeneous combination of symptom clusters that vary in presentation, duration, and severity from child to child (Cicchetti & Toth, 1998). Essentially, two different children can have the same diagnosis, but present with different manifestations of the disorder. For example, in order to meet the DSM-IV (APA, 1994) criteria for a diagnosis of major depressive episode, an individual must present with a minimum of five out of nine criteria (i.e., symptoms). It is conceivable that two children could be diagnosed with only a single common symptom. Therefore, both children can be diagnosed appropriately with major depressive episode; however, their manifestations of major depressive episode are quite different. Subsuming children with different symptom presentations into a broader diagnostic category leads to an inevitable loss of data regarding the true variability and unique symptom constellations presented across children suffering from depressive sequelae. Subsequently, treatment
decisions might be based on broad assumptions about major depressive episode without considering the specific needs of individual children.

A third factor that complicates the assessment and diagnosis of depression in children is the issue of comorbidity. According to Stedman’s Medical Dictionary (1995) comorbidity is defined as “a concomitant but unrelated pathological or disease process” (p. 174). For example, an individual might be suffering from both lung cancer and Hepatitis B at the same time and thus be considered to have comorbid medical illnesses. These two disease processes are essentially independent of one another, with different etiologies, symptom presentations, and progression patterns.

While the medical definition of comorbidity implies that illnesses are concomitant but unrelated, the use of the term comorbidity in the psychological and psychiatric literature is less well-defined. Unlike many medical illnesses, psychological disturbances are not discrete illnesses and are therefore more difficult to assess, diagnose, and classify, due, in part, to the overlapping nature of the various symptom clusters (Adams & Cassidy, 1993). For example, two disorders in the DSM-IV (APA, 1994), major depressive episode and generalized anxiety disorder, have overlapping diagnostic criteria including: irritability, difficulty concentrating, sleep disturbance, and fatigue (see Figure 1). Consequently, children who present with these symptoms would meet some of the diagnostic criteria for MDE while simultaneously satisfying some of the diagnostic parameters of generalized anxiety disorder (GAD), thus making an accurate differentiation between the two diagnoses problematic. Angold and Costello (1993) stated that researchers and clinicians must evaluate whether patterns of comorbidity are
Figure 1. DSM-IV (APA, 1994) symptom overlap: major depressive episode and generalized anxiety disorder.

"artifacts of the methods of data collection, data aggregation for diagnostic purposes, or the nosology itself" (p. 1786). In sum, various authors have suggested that assessment methods should be broad enough (e.g., multiple instruments) to accurately identify several different constellations of symptoms, especially when there is evidence of comorbidity (Achenbach, 1985; Costello, 1986; Reynolds, 1992a, 1992b).

In spite of these diagnostic limitations, there is mounting evidence that the existence of more than one psychological disorder in the same child is prevalent and can lead to poorer outcomes. For example, Offord et al. (1992) followed a community sample of over 1,000 children between 4 and 12 years old and reported that children with
multiple disorders tend to “have a worse outcome years later compared with the single-
disorder group” (p. 921). In a more recent longitudinal study involving over 1,000
children who were followed from birth to age 21, Newman et al. (1996) reported that
nearly half of the subjects who evidenced a psychiatric disorder during the course of the
study also had comorbid diagnoses at the age of 21. Further, the authors reported that
“comorbidity was associated with severity of impairment” (p. 552).

Current comorbidity estimates for depression and anxiety in children range from
15.9% to 61.9% (Brady & Kendall, 1992). Anderson et al. (1987) examined a
nonclinical sample of 63 children and found that 15.9% qualified for both an anxiety
disorder and a depressive disorder. However, Costello et al. (1988) reported much lower
estimates of comorbidity in a nonclinical group of pediatric primary care patients, with
coexisting symptoms of depression and anxiety appearing in 0.8% of the sample. In
clinical samples, the comorbidity rates have been much higher. In a sample of
hospitalized children, Carey, Finch, and Imm (1989) reported that 55.2% of the sample
had diagnosable disorders of both depression and anxiety. Furthermore, Strauss, Last,
Hersen, and Kazdin (1988) reported that in a group of outpatient children and adolescents
who presented with anxiety disorders, 28.3% of the sample also met criteria for
depressive disorders.

Based on the aforementioned findings, there is a great deal of variation in the
comorbidity estimates of depression and anxiety in children and adolescents. The
variation has been attributed to “rather crude diagnostic criteria” (Angold & Costello,
1993, p. 1786), unreliable data collection techniques (Garfield, 1993), similarity in self-
report rather than construct overlap (Norvell, Brophy, & Finch, 1985), and the fact that depression and anxiety might be clinically related (Clark, Beck, & Stewart, 1990; Ollendick & King, 1994). Despite the discrepant findings, several authors have suggested that the overall comorbidity rates are large enough to be considered clinically meaningful (Kendall, Kortlander, Chansky, & Brady, 1992; Newman et al., 1996; Reynolds, 1992b).

In sum, the assessment and diagnosis of depression in children is a complex process with a number of limitations. Nonetheless, it is incumbent upon researchers and clinicians to understand this process and to use the best methods available in order to best serve these children. In light of the variation in comorbidity estimates and the diverse presentation of symptom constellations, various researchers have recommended assessment practices that emphasize broad-band instruments and the solicitation of information from several sources (Achenbach et al., 1987; Finch & Rogers, 1984; Kazdin, 1988; Reynolds, 1992c). Similarly, interventions may be differentially effective for various children (i.e., comorbid vs. no comorbid diagnoses); thus clinicians and researchers must remain mindful of these considerations when evaluating the efficacy of treatments.

Summary

Although childhood depressive disorders were not recognized until recently, it is now known that depression in children is a serious problem that has critical ramifications for a large number of young people. Depressive disorders typically result in a great deal of distress and suffering in children, which, for some, can lead ultimately to fatal
consequences in the case of suicide. In light of the prevalence, persistence, and serious consequences associated with childhood depression, mental health professionals have an obligation to provide expeditious and effective treatments to young people suffering from depressive disorders.

Treatment of Childhood Depression

Several treatments have been developed for depression in children ranging from group-based cognitive-behavioral interventions to pharmacotherapies. While there are a growing number of well-controlled clinical trials, most of the early attempts to document the efficacy of treatments for child and adolescent depression were based upon clinical experience, case reports, and extrapolations from the adult literature (Mueller & Orvaschel, 1997; Speier, Sherak, Hirsch, & Cantwell, 1995). In contrast, much more is known about the treatment of adult depressive disorders since a large number of well-controlled clinical trials have been conducted, which in turn, have been integrated and summarized via meta-analytic studies (e.g., Robinson, Berman, & Neimeyer, 1990). In light of the paucity of clinical trials for child and adolescent depression relative to adult samples and the lack of any published, methodologically sound integrative reviews, the current state of knowledge regarding the efficacious treatment of childhood depression is unclear. This section includes a review of the adult and child literatures regarding the treatment of depression.

Adult Literature

A number of treatments for depression have been empirically studied and
validated for adult populations. For instance, the National Institute of Mental Health (NIMH) sponsored a large, multisite, controlled clinical trial that compared the efficacy of four treatments for adult depression, including imipramine plus case management (IMI-CM), a placebo plus case management (PLA-CM), and two types of psychotherapy (Treatment of Depression Collaborative Research Program; Elkin, 1994). The two types of psychotherapy implemented were cognitive-behavioral therapy (CBT) and interpersonal therapy (IPT). While there is some variation in the components of CBT, the ultimate goal of this therapeutic approach is to identify irrational and dysfunctional thoughts, become aware of how they lead to depressive feelings and behavior, and to modify thinking patterns (Beck, Rush, Shaw, & Emery, 1979). IPT, on the other hand, focuses on dysfunctional patterns of relating to others in a variety of social milieus with the main goal of identifying and modifying problematic interpersonal tendencies (Elkin, Parloff, Hadley, & Autry, 1985). In general, the findings indicated that either type of psychotherapy was as effective as medication and more effective than PLA-CM in reducing depressive symptomatology in a group of 250 outpatient adults. Elkin (1994) reported that "there were...no statistically significant differences between the two psychotherapies or between either of them and IMI-CM" (p. 119).

Furthermore, there have been a number of large integrative reviews investigating the efficacy of various treatments for adult depressive disorders. For example, Robinson et al. (1990) conducted a comprehensive review of 58 controlled outcome studies of various psychotherapeutic treatments for adult depression (e.g., cognitive-behavioral, psychodynamic, client-centered, pharmacologic). Based on their sample of studies, the
authors reported a standard mean difference effect size (ES; mean of the treatment group minus the mean of the control group divided by the control group or pooled standard deviation) of 0.74 when comparing psychotherapy clients and no treatment controls at posttreatment. This finding indicates that the average treated client was better off at posttreatment than approximately 77% of the untreated clients. Robinson et al. concluded that depressed adults derive substantial and long-lasting benefits from psychotherapy. They added that the benefits of psychotherapy were at least as effective as medication and that no specific type of psychotherapy was superior to another in assuaging depression in adult samples. Thus, important questions remain regarding what treatments are best for whom even with adults.

Despite the fact that there are empirically supported methods of treating adult forms of depression, the generalization of these findings to children might be problematic in light of the developmental, cognitive, emotional, and age differences observed between adults and children (Speier et al., 1995). Yet, to a large extent, the adult literature has been used to guide the efforts to treat child and adolescent depression (Holmes & Wagner, 1992; Mueller & Orvaschel, 1997). While this extrapolation approach makes intuitive sense, making inferences from the adult literature to the child literature requires a leap of faith, a leap that has not been empirically validated to date.

Child Literature

While the psychotherapy literature for children and adolescents is not nearly as substantial as the adult literature, there have been a large number of treatment studies for
youth who present with a broad array of problems. Further, the results from several studies have been synthesized in large meta-analytic reviews of the overall efficacy of psychotherapy with children and adolescents. For example, Casey and Berman (1985) reviewed 75 studies published between 1952 and 1983 and reported a standard mean difference ES of 0.71, indicating that the average treated child scored better after treatment than approximately 76% of control-group children. However, there were no studies that specifically addressed depression in children. In another large meta-analytic review, Weisz, Weiss, Alicke, and Klotz (1987) examined the outcomes of 108 well-designed treatment studies published between 1952 and 1983 involving children between the ages of 4 and 18 years of age. The authors reported a standard mean difference ES of 0.79, indicating that the average child who received treatment was at the 79th percentile of control-group children. Kazdin, Bass, Ayers, and Rogers (1990) surveyed the published literature between 1970 and 1988 and included a broad sample of 223 studies in their analysis. When comparing treated versus untreated controls (64 studies), Kazdin et al. reported a standard mean difference ES of 0.88, indicating that the average child in the treatment group was better off at posttreatment than 81% of the control-group children. However, the investigators in all of the aforementioned reviews did not specifically address the issue of treating child and adolescent depressive disorders or they collapsed “depression” studies into more broadly defined categories (e.g., internalizing, emotional problems), making the extraction of specific data regarding the efficacy of depression treatment problematic. For example, Kazdin et al. (1990) categorized 16% of the original sample of 223 studies as treatment trials for “internalizing” problems (e.g.,
anxiety, withdrawal, emotional disturbances, depression).

Only one large published integrative review was located that addressed the issue of treating child and adolescent depression. In a meta-analytic review of 150 outcome studies on child and adolescent psychotherapy, Weisz et al. (1995) included six controlled outcome studies of child and adolescent depression and reported a standard mean difference ES of 0.64, indicating that the average child who received treatment for depression was better off than approximately 74% of children who did not. In a more recent investigation, Reinecke, Ryan, and DuBois (1998) analyzed many of the same studies (i.e., four of the six) included by Weisz, Weiss, et al. (1995) and included two others (i.e., Lerner & Clum, 1990; Wood, Harrington, & Moore, 1996) in a highly focused meta-analytic review of controlled cognitive-behavioral interventions for depressed adolescents. Reinecke et al. (1998) reported a standard mean difference ES of 1.02, indicating that the average treated adolescent was better off at posttreatment than approximately 84% of the control subjects. Although the findings from these investigations are important, Weisz, Weiss et al. (1995) included only six studies and they did not specifically address how the outcomes of child depression studies might covary or interact with other variables (e.g., age, sex, type of treatment). In terms of the meta-analytic review by Reinecke et al. (1998), the sample of studies analyzed were not only similar to a previously published meta-analysis, but the findings were limited in the sense that only cognitive-behavioral interventions for adolescents were included. Thus, the findings were far from complete. Essentially, one is left to ponder several important questions regarding the efficacy of psychosocial treatments for child and adolescent
depression. For example, what treatments are the most effective and for whom? Does the age of the child make a difference? Do other variables such as length of treatment, therapist variables, and comorbid conditions make a difference? Is there differential efficacy for those suffering from depressive symptomatology versus a diagnosed depressive disorder? Furthermore, how do psychosocial interventions compare with pharmacological treatments? In sum, is there evidence of differential efficacy when a number of important variables are considered such as age, sex, severity, and methodological factors?

In one other systematic review worthy of mention, Black-Cecchini (1996) analyzed 13 outcome studies investigating the efficacy of group-based social skills interventions for depressed youth as part of an unpublished doctoral dissertation. By using meta-analytic procedures, Black-Cecchini reported that the social skills group interventions produced moderate to large effects (i.e., average ES 0.76) in terms of ameliorating depressive symptomatology in depressed children and adolescents primarily from school samples. Nonetheless, Black-Cecchini (1996) suggested that some of the outcomes covaried with certain design and sample characteristics. For example, Black-Cecchini cross-tabulated the quality of the studies with the outcomes and found that “inferior” studies had the highest mean effect sizes when compared to the higher quality studies. In addition, Black-Cecchini noted that it was difficult to determine how other variables (age, ethnic status, type of treatment) covaried with the outcomes because either the data were not reported or the data were poorly defined or quantified. Further, Black-Cecchini described some of the limitations found in the sample of outcome studies.
including a paucity of investigations which included preadolescent samples (7-12 age group; N = 4) and an overreliance on self-report measures. In essence, Black-Cecchini’s analysis provided some provocative ideas about the nature and course of the present study.

Psychosocial Treatments for Childhood Depression

As mentioned previously, a variety of psychosocial treatment regimens for depressed children have been described in the literature including: psychoanalytic, behavioral, familial, group, cognitive, relaxation, social skills training, problem-solving training, and multimodal therapies. For instance, in one case study, Frame, Matson, Sonis, Fialkor, and Kazdin (1982) used behavior therapy to treat a 10-year-old depressed child and reported that certain behavioral indices associated with depression (e.g., suicidal gestures, labored psychomotor functioning) improved during the course of treatment and after a 12-week follow-up. Other investigators have reported similar results when using behavioral regimens to treat depressed youth (e.g., Matson, 1982).

In addition, some of the behavioral modalities are downward extensions of interventions originally intended for adults. For example, a behavioral method of treating depression in adults is based on the premise that depression results from low rates of environmental reinforcement, which is manifested by diminished activity levels (Costello, 1972; Lewinsohn, 1974). Interventions designed to increase activity levels, especially pleasant activities, purportedly led to an increase in the amount of response-contingent reinforcement for clients, resulting in a decrease in depressive symptomatology (Kaslow
& Rehm, 1991). Furthermore, there is some evidence to support the efficacy of activity-based regimens in adult depressives (e.g., Rehm & Kaslow, 1984).

The relationship between pleasant/unpleasant activities and depression in youth has been examined empirically. In a sample of 8- to 14-year-old children, Wierzbicki and Saylor (1991) reported a positive correlation between elevated symptoms of depression and unpleasant activity levels. However, contrary to the authors’ predictions, they did not find a negative correlation between self-reported depressive symptomatology and pleasant activity levels. Although there have not been any depression treatment studies in child samples that adhere strictly to an activity-increase regimen, several empirical investigations have used activity-based interventions as part of the treatment protocol (e.g., Clarke et al., 1995; Kahn, Kehle, Jenson, & Clark, 1990; Lewinsohn, Clarke, Hops, & Andrews, 1990; Stark, Reynolds, & Kaslow, 1987). As such, it remains unclear as to whether these activity-based components contribute substantially to treatment efficacy.

Similar to activity-increase regimens that are based on the idea that diminished reinforcement leads to depressive symptomatology, social skills models are premised upon the belief that underdeveloped or impaired social skills often preclude individuals from obtaining adequate social reinforcement that subsequently leads to depression (Rehm & Kaslow, 1984). Social skills training (SST) has been used to treat depression in children and adolescents. For example, Fine, Forth, Gilbert, and Haley (1991) compared a social skills training group and a therapeutic support group in the treatment of depression in adolescents. The authors reported that the therapeutic support group was more effective than the social skills group in ameliorating depressive symptoms in the
adolescent sample at posttreatment, but that the modalities appeared to be equally effective after a 9-month follow-up.

In another study, King and Kirschenbaum (1990) reported that social skills training plus consultation was superior to consultation only in the prevention of depressive symptoms in a cohort of “at-risk” children between 5 and 11 years old. Furthermore, a number of other treatment studies involving depressed children and adolescents have included SST components in the intervention protocols (e.g., Clarke et al., 1995; Kahn et al., 1990; Lewinsohn et al., 1990; Liddle & Spence, 1990; Reynolds & Coats, 1986; Stark et al., 1987). However, questions remain about the degree to which social skills training components impact treatment outcomes.

Other types of psychosocial interventions have been used in the treatment of child and adolescent depression including: aerobic exercise (Brown, Welsh, Labbé, Vitulli, & Kulkarni, 1992), relaxation (Kahn et al., 1990; Reynolds & Coats, 1986; Wood et al., 1996), psychoeducation (Clarke, Hawkins, Murphy, & Sheeber, 1993), and a combination of group and family therapy (Curry & Wells, 1998). The results from these studies have been mixed, with substantial differences in methodology, selection, and treatment procedures. Furthermore, Mufson et al. (1994) reported that a variant of IPT was efficacious in the treatment of early-onset major depressive disorder in a cohort of depressed adolescents.

Perhaps the most common method of treating depressed youth in empirical trials is CBT, with the majority of studies favoring group over individual interventions. As mentioned above, cognitive-behavioral treatments are aimed at identifying irrational and
dysfunctional thoughts, enhancing awareness of how such thoughts lead to depressive feelings and behavior, and modifying thinking and behavioral patterns (Beck et al., 1979). In what is considered by many to be the seminal empirical study on the treatment of depressed youth, Butler, Miezitis, Friedman, and Cole (1980) compared role-play therapy, cognitive restructuring, and a wait-list control condition in a school-based sample of 56 depressed children. The authors reported that both treatments were more effective in reducing depressive symptomatology than the wait-list control condition, with role-play therapy being slightly superior to cognitive restructuring.

In another investigation, Reynolds and Coats (1986) conducted a group treatment study of child and adolescent depression with 30 moderately to severely depressed high school students. The investigators compared three treatment conditions including CBT, relaxation training, and a wait-list control group and reported that both active treatments resulted in significant improvements in depressive symptomatology when compared to wait-list controls at posttreatment and after a 5-week follow-up (Reynolds & Coats, 1986).

Furthermore, Weisz, Thurber, Sweeney, Proffitt, and LeGagnoux (1997) implemented a group-based, eight-session Primary and Secondary Control Enhancement Training program (see Rothbaum, Weisz, & Snyder, 1982 for further review) with 48 children with mild-to-moderate depressive symptoms. Weisz et al. (1997) reported that the intervention was effective in reducing self-reported symptomatology at posttreatment and after a 9-month follow-up in comparison to control subjects. In addition, there is a recent trend of examining the efficacy of individual CBT in the treatment of depressed
youth with some promising results (e.g., Brent et al., 1997; Vostanis, Feehan, Grattan, & Bickerton, 1996; Wood et al., 1996).

While the aforementioned studies provide us with important insights about the treatment of childhood depression, many questions are left unanswered. Based upon a general review of the literature, it is estimated that there are approximately 20-25 empirical studies investigating the efficacy of psychosocial treatments for child and adolescent depression to date. However, there are important differences across the various studies (e.g., design, selection, outcome, number of subjects, setting, assessment/diagnostic strategies, therapist variables, age range, demographics, type, length, frequency, and intensity of treatment). Thus, useful generalizations about which interventions would be the most effective and for whom are difficult to make based on a general review of the literature. These findings highlight the need to integrate the body of literature on the psychosocial treatments for childhood depression.

Narrative Reviews

Thus far, most of the efforts to review, integrate, and summarize the results of various psychosocial treatment outcome studies for child and adolescent depression have been inadequate for guiding research, policy, and practice because many of the reviews are narrative in nature, methodologically flawed, and/or present vague or conflicting conclusions. Thirteen narrative reviews were located, published between 1983 and 1998 (Birmaher, Ryan, Williamson, Brent, & Kaufman, 1996a, 1996b; Clarizio, 1986; Cytryn & McKnew, 1985; Dujovne, 1993; Harrington, 1992; Holmes & Wagner, 1992; Kashani
& Cantwell, 1983; Kaslow & Thompson, 1998; Keller, Lavori, Beardslee, Wunder, & Ryan, 1991; Larsson, 1992; Petersen et al., 1993; Puig-Antich & Weston, 1983; Reynolds, 1990). The number of psychosocial treatment studies cited in the above mentioned reviews ranged from zero (Cytryn & McKnew, 1985) to 14 (Kaslow & Thompson, 1998). In terms of the substantive findings from the earlier reviews, they are often unclear and conflicting. For example, Clarizio (1986) provided a narrative review of the treatments for childhood depression, including behavioral regimens, cognitive strategies, pharmacotherapy, family therapy, and multimodal therapy. In general, Clarizio suggested that there is evidence that many of these regimens are effective in ameliorating depressive symptomatology in children. However, Clarizio (1986) cited only s treatment studies in his review, he did not include any effect sizes, and he cautioned that “there are no easy answers” when choosing a treatment method for childhood depression.

In another narrative review, Larsson (1992) concluded that “controlled investigations of psychological treatments have, to date, been carried out solely with school samples of children that suffer from moderate depression” (p. 12). Larsson cited seven treatment studies and he added that most of these treatments were based on cognitive-behavioral principles (e.g., problem-solving, cognitive-restructuring). In addition, Larsson reported that despite the evidence that cognitive-behavioral interventions have been shown to be efficacious with school samples, investigations on the effectiveness of psychotherapy for depressed children from clinical populations are nonexistent.

In more recent reviews of the treatments for child and adolescent depression, the
conclusions are neither compelling nor are they based on exhaustive surveys of the literature. Dujovne, Barnard, and Rapoff (1995) included approximately treatment studies in their review and concluded that “the efficacy of cognitive-behavioral treatments has not yet been consistently demonstrated because investigations have been sparse and lacking in methods of subject selection and assessment which could result in representative samples” (p. 606). Approximately one year later, Birmaher et al. (1996a) reviewed the previous years of literature on child and adolescent depression. While the authors cited more psychological treatments studies (i.e., 10) than previous reviewers, the survey was far from comprehensive and the authors did not systematically analyze how treatment outcomes might covary with important variables (e.g., age, sex, severity of illness, treatment type).

In perhaps the most systematic narrative review to date, Kaslow and Thompson (1998) reviewed 14 psychosocial interventions for child and adolescent depression. In general, the authors suggested that “psychosocial interventions are effective at posttreatment and follow-up in reducing depressive symptoms/disorders in clinical and nonclinical samples of youth, regardless of treatment modality or extent of parental involvement” (p. 146). However, Kaslow and Thompson (1998) examined the extent to which each of the studies conformed to guidelines proposed by the Task Force on Promotion and Dissemination of Psychological Procedures (1995) “for well-established and probably efficacious interventions” (p. 146). Briefly, the criteria for well-established psychosocial interventions for childhood disorders are: (a) at least two well-conducted group-design studies conducted by different research teams indicate the treatment is
superior to a pill placebo or an alternative treatment, or equivalent to an already
established treatment; or (b) a large series of methodologically sound single case studies
(i.e., \( n > 9 \)); (c) interventions where treatment manuals were utilized; and (d) clearly
specified sample characteristics (Lonigan, Elbert, & Johnson, 1998). The criteria for
probably efficacious psychosocial interventions for childhood disorders are slightly less
stringent in that they do not require replication by two different research teams, the
treatment must be more effective for experimental versus no-treatment control subjects,
and the required number of methodologically sound single case designs was reduced (i.e.,
\( n > 3 \)). In sum, Kaslow and Thompson (1998) concluded that only 2 of the 14 studies
reviewed met the criteria for probably efficacious psychosocial interventions for
childhood disorders (i.e., Stark et al., 1987; Stark, Rouse, & Livingston, 1991). In
addition, the authors reported that not one of the reviewed studies met criteria for well-
established psychosocial interventions for childhood disorders. Thus, even the most
recent reviews of the child and adolescent depression treatment literature leave many
questions unanswered and unaddressed.

In summary, published narrative reviews of the child and adolescent depression
treatment literature are plentiful yet uncompelling in guiding policy, research, and
practice. Further, in the reviews published to date, there has not been an attempt to
collect and systematically compare a more comprehensive sample of treatment studies,
including pharmacologic interventions. Thus, a more complete, systematic analysis of the
child and adolescent depression treatment literature is clearly indicated.
Pharmacological Treatment

The prescription of pharmacologic medications for depression in children has become a widely accepted practice (Kaplan, Simms, & Busner, 1994) despite the fact that the Food and Drug Administration (1994; FDA) has yet to approve a single psychotropic agent for the specific purpose of treating depressed youth up to the age of 17 years old (Physician’s Desk Reference, 1998; Peter Jensen, personal communication, March 16, 1998; Sommers-Flanagan & Sommers-Flanagan, 1996). The practice of using approved drugs for unlabeled conditions is common and does not prohibit physicians from such practices (American Academy of Pediatrics Committee on Drugs, 1996). Unapproved use of approved drugs does not necessarily imply disapproval or contraindication, only that the drug has not been studied adequately for a particular (i.e., unlabeled) condition. Before a drug can be approved for a specific problem, its safety and efficacy must be evaluated and the data submitted to the FDA. Realistically, this process can take several years, and thus physicians are often faced with challenge of balancing their clinical judgment with the available empirical literature at the time. The labeling of drugs “is intended neither to preclude the physician from using his or her best medical judgment in the interest of patients nor to impose liability for failure to comply with labeling restrictions” (American Academy of Pediatrics, 1996, p. 144).

In cases of “off label” prescribing, Spencer, Wilens, and Biederman (1995) suggested that the “risks, potential benefits, and informed consent should be carefully documented” (p. 97). Thus, a drug’s efficacy and its potential side effects must be carefully evaluated and reviewed with patients and their families before initiating a
medication trial. In the case of pharmacotherapy for child and adolescent depressive disorders, psychiatrists and pediatricians can choose from a variety of psychotropic medications including tricyclic antidepressants (TCAs) such as imipramine (Tofranil), amitriptyline (Elavil), desipramine (Norpramin), and selective serotonin reuptake inhibitors (SSRIs) including fluoxetine (Prozac), paroxetine (Paxil), and sertraline (Zoloft). Other medications that have been prescribed for children and adolescents who present with atypical depression, treatment-resistant mood disorders, and bipolar forms of affective illness include monoamine oxidase inhibitors, mood stabilizers such as lithium or valproic acid, and anticonvulsants such as carbamezapine (Spencer et al., 1995). However, the types of psychotropic medications used most often for children who present with depression are TCAs and SSRIs (Birmaher et al. 1996).

It has been hypothesized that medications such as TCAs and SSRIs act on the noradrenergic and serotonergic systems in the central nervous system by slowing the reuptake of particular neurotransmitters (e.g., norepinephrine, serotonin) in the synaptic cleft, thus prolonging postsynaptic potentials (Kandel, 1991). Although several hypotheses regarding the action of antidepressant medications have been advanced over the past 30 years (e.g., catecholamine hypothesis), a compelling model of the exact nature of the pharmacologic mechanisms and the pathophysiology of depression remains elusive (Duman, Heninger, & Nestler, 1997). Recent empirical findings suggest that antidepressant medications impact brain physiology well beyond the neurotransmitter (i.e., monoamine) and receptor levels. For example, Duman et al. (1997) suggested that the “long-term, therapeutic action of antidepressant treatments is mediated by
postreceptor intracellular targets” (p. 598; see Duman et al., 1997 for a review).

With regard to the safety of pharmacotherapy for child and adolescent depression, the potential adverse impact that pharmacologic interventions have on child development remains unclear (Antonuccio, Danton, & DeNelsky, 1995; Vitiello & Jensen, 1997). Further, there have been some serious concerns raised about treating children with TCAs. For example, Riddle, Geller, and Ryan (1993) reported that the sudden death of a child was associated with the cardiotoxic effects of the antidepressant medication desipramine. In addition, Werry (1995) seriously challenged the safety of using certain TCAs with children in light of their noxious side effects (e.g., anticholinergic effects, sedation), and the increasing number of deaths associated with the use of certain TCAs in children. However, some authors have suggested that the association between TCAs and cardiotoxic effects in children remains unclear and should not necessarily preclude their use in juvenile patients (Wilens et al., 1996). With the advent of SSRIs, which are known to have less noxious side effect profiles (e.g., limited anticholinergic and sedative effects) and are easier to administer (i.e., once per day) than most of the TCAs, using SSRIs might be a safer alternative to the use of TCAs for the treatment of child depressive disorders (Birmaher et al., 1996a).

As mentioned previously, the practice of child and adolescent pharmacotherapy for depression makes sense in light of the strong parallels (i.e., syndromally and diagnostically) between child and adult manifestations of depressive sequelae (Ryan et al., 1987) and the empirical evidence that the use of psychotropic medications in adults suffering from depression is efficacious, with an average positive response rate between
65% and 75% (Speier et al., 1995; Thase & Kupfer, 1996). In a recent revision of the FDA regulations (1994) regarding the establishment of efficacy for “pediatric use” of pharmacologic agents, it was noted that extrapolations of drug efficacy from adults to children are justified provided that there is compelling evidence that “the course of the disease and the drug’s effects are sufficiently similar in the pediatric and adult populations” (p. 64241). While there is evidence of psychopathological congruity between child and adult depressive disorders, Vitiello and Jensen (1997) cautioned that “similar treatment response cannot be inferred across the age groups based only on similar psychopathological or biological features” (p. 872). The authors added that despite the similarities in depressive phenomenology across the age groups, a “discrepancy remains between the efficacy of tricyclic antidepressants in adults versus that in children” (p. 872, Vitiello & Jensen, 1997).

Despite the equivocal status of treating child and adolescent depression with TCAs, early studies on the efficacy of antidepressant medication in child and adolescent samples appeared to be promising. Weinberg et al. initiated an open medication trial for 34 depressed children, using either one of two TCAs, imipramine or amitriptyline (Weinberg, Rutman, Sullivan, Penick, & Dietz, 1973). The authors reported that 12 of the 19 children who received the medication evidenced “marked improvement” after one month. In contrast, it was reported that only 3 of the 15 children who did not receive the antidepressant medication evidenced clinical improvement after one month (Weinberg et al., 1973). In another open trial, Puig-Antich, Blau, Marx, Greenhill, and Chambers, (1978) reported that six out of eight children evidenced a substantial reduction in
depressive symptomatology after eight weeks of treatment on imipramine.

Other open medication trials appeared to be effective in ameliorating depressive symptoms in children as well (e.g., Geller, Cooper, Chestnut, Anker, & Shuchter, 1986; Preskorn, Weller, & Weller, 1982). However, as mentioned, these studies were open trials without control groups, double-blind conditions, or pharmacologically inert placebos. Thus, the positive findings were questioned on empirical grounds. That is, it was difficult to determine whether the observed improvements in the children were due to patient expectations (i.e., "placebo-effects"), measurement error, researcher bias, the mere passage of time, or other artifactual confounds. Placebo-controlled trials are of particular interest since a placebo effect is defined as "a beneficial effect... that arises from the patient's expectations concerning the treatment rather than from the treatment itself" (Stedman's Medical Dictionary, 1995, p. 644). In antidepressant trials for adult samples, Morris and Beck (1974) reported that the placebo response rates ranged between 30% and 35%, which is generally considered to be an average placebo effect for antidepressants (Burke & Preskorn, 1995).

In light of the aforementioned methodological criticisms of open studies, a number of randomized, placebo-controlled medication trials for child and adolescent depression were designed and implemented. For example, Puig-Antich et al. initiated a placebo-controlled imipramine trial in a sample of 42 depressed children between the ages of 6- and 12-years-old (Puig-Antich et al., 1978). The authors reported a higher response in the placebo condition (68%) than in the active medication condition (56%). These findings were somewhat troubling since they contradicted the previous "evidence"
of antidepressant efficacy in children reported from the open trials. Kye and Ryan (1995) reviewed nine randomized, double-blind, placebo-controlled antidepressant trials in child and adolescent samples. The authors reported that only one study (i.e., Preskorn, Weller, Hughes, Weller, & Bolte, 1987) provided evidence that pharmacotherapy was superior to placebo in ameliorating depressive symptomatology in a group of 22 children between the ages of 6 and 12 years old. Further, Kye and Ryan (1995) cautioned that the evidence of efficacy in the study by Preskorn et al. was “clinically small but statistically significant” (p. 268). Preskorn et al. (1987) reported that the improvement from baseline was 43% in the active medication condition and 35% in the placebo condition.

Only one published meta-analytic study was located in which the placebo-controlled, double-blind trials using TCAs in the treatment of depression in children and adolescents were reviewed. Hazell, O’Connell, Heathcote, Robertson, and Henry (1995) reviewed 12 randomized, controlled trials comparing the efficacy of TCAs and placebo in depressed children between 6 and 18 years of age. The authors reported a standard mean difference ES of 0.35 for the active medication conditions versus the placebo (control) conditions. Hazell et al. (1995) concluded that “the small additional effect afforded by treatment in comparison with placebo is unlikely to be clinically important in most patients” (p. 899). However, according to the authors, they were only able to derive effect sizes from 6 of the 12 studies due to limited information contained in the articles. Thus, some authors have criticized the conclusions from this meta-analytic review on empirical grounds (e.g., Anderson, 1995). Furthermore, in a review of the drug trials for early-onset depression, Conners (1992) cautioned that “various methodological problems
limit the conclusions that can be drawn" regarding the efficacy of antidepressant medications (p. 11). That is, the substantial differences in the measurements used, criteria for “improvement,” selection procedures, and sample characteristics in the various studies render comparisons across the studies tenuous at best (Conners, 1992).

Although the literature regarding the use of TCAs in the treatment of child and adolescent depression has been either discouraging or equivocal, there have been several recent studies involving the use of SSRIs with promising results. In a retrospective review of the medical records of 31 children between 9 and 18 years old who were hospitalized for depression, Jain, Birmaher, Garcia, Al-Shabbout, and Ryan (1992) reported that after a mean treatment duration of 35 days on fluoxetine, clinical improvement was seen in 74% of the patients as measured by the Clinical Global Impression Scale. In another retrospective review, Tierney, Joshi, Llinas, and Rosenberg (1995) evaluated the efficacy of sertraline in a group of 33 children and adolescents between the ages of 8 and 18 years old. The authors reported that the majority of the children evidenced significant improvement based on the indices of Clinical Global Impression Scales after the 10th week of treatment (Tierney et al., 1995). Moreover, open trials with fluoxetine have reportedly produced a significant reduction of depressive symptomatology in child and adolescent patients as well (e.g., Boulos, Kutcher, Gardner, & Young, 1992; Collie, Belair, Difeo, Weiss, & LaRoache, 1994).

In addition to retrospective reviews and open trials, the results from several controlled studies involving the use of SSRIs in child and adolescent samples have been published recently. Simeon, Dinicola, Ferguson, and Copping (1990) conducted a study
in which 40 adolescents between 13 and 18 years old participated in a placebo-controlled, double-blind study of fluoxetine treatment for depression. The authors reported that "fluoxetine was superior to placebo on all clinical measures except for sleep disorder, but the differences were not statistically significant" (Simeon et al., 1990, p. 71). In another trial, Mandoki, Tapia, Tapia, and Sumner (1997) evaluated the efficacy of venlafaxine (Effexor) in a cohort of 33 children and adolescents who met the DSM-IV (APA, 1994) criteria for major depressive disorder. The two treatment conditions consisted of venlafaxine and psychotherapy or placebo and psychotherapy. The authors reported that significant improvements were observed in both conditions, but they were reluctant to attribute the positive effects to venlafaxine drug therapy (Mandoki et al., 1997). However, in a recently published controlled trial of 96 children and adolescents being treated with fluoxetine, Emslie et al. (1997) reported a positive response rate of 56% in the active condition compared to a 33% positive response rate for those taking placebo. Thus, it appears that there have been some recent positive findings regarding the efficacy of SSRIs in children and adolescents with depressive disorders.

Overall, it is estimated that there are between 14-17 published studies investigating the efficacy of psychotropic medications for depressed youth. A more up-to-date comprehensive review that integrates the recent studies and includes a systematic analysis of some of the important variables associated with outcome (e.g., length of treatment, outcome measures) might provide some answers to important questions related to the efficacy of pharmacotherapy for child and adolescent depression. Furthermore, there is a paucity of data on the comparison of pharmacological and psychosocial
interventions for depressed children and adolescents. Only two studies were located that compared psychotherapy versus pharmacotherapy in the treatment of early-onset depression. In one of the studies (Mandoki et al., 1997, cited above), a sample of adolescents with major depression were treated with either venlafaxine and psychotherapy or pill placebo and psychotherapy. Although both treatments appeared to ameliorate depressive symptoms to a substantial degree, Mandoki et al. (1997) reported that “there was not a significant difference in improvement between both groups, which indicates that the medication treatment did not contribute significantly to the improvement” (p. 153). In the only other study located that compared psychosocial and pharmacological interventions, Dujovne (1993) treated six clinically depressed children between the ages of 8 and 11 with imipramine and cognitive-behavioral therapy in a multiple-baseline, crossover design. Dujovne (1993) reported that both treatments resulted in a significant reduction in depressive symptoms over the course of the study with a slight advantage to the cognitive-behavioral regimen and increased parental satisfaction of the nonpharmacologic intervention (i.e., cognitive-behavioral therapy). However, the author cautioned that four of the six children evidenced noxious side-effects while taking the imipramine and the medication was discontinued. In sum, it appears that there are some important questions to be answered regarding the comparison of psychosocial and pharmacological interventions for depressed youth. The most compelling question seems to be whether pharmacotherapy should be the first-line treatment for children and adolescents who present with depression when nonpharmacologic interventions may be just as effective and do not expose the patient to
potentially noxious, if not dangerous side-effects.

Summary

In summary, a number of treatments for childhood depressive disorders have been developed ranging from pharmacotherapy to psychosocial regimens and it is estimated that there are approximately 40-50 studies investigating the efficacy of interventions for depressed youth. What we do know about the efficacy of psychosocial treatments for childhood depression is limited. There is moderate support for the efficacy of school-based group interventions for children and adolescents who evidence depressive symptomatology on self-report measures. However, beyond this, there is not much else that can be said about empirically validated psychosocial interventions for childhood depression. In terms of pharmacologic interventions, it appears that TCAs have not been found to be superior to placebo in the majority of controlled outcome studies. There have been some recent studies of the efficacy of TCAs and SSRIs in the treatment of children with depression, but these findings have yet to be integrated into a comprehensive literature review. Although some of the efforts to review and synthesize the body of literature on the treatments for depressed youth have answered some questions, there are many more questions left unanswered. For instance, which variables need to be considered? Do outcomes covary with study and sample characteristics? Which types of treatments are the best and for whom? Does the length of treatment impact outcomes? Does the age of the child matter? Do comorbid conditions impact the outcomes? Does it make a difference whether the children have been identified as suffering from depressive
symptomatology as opposed to a depressive disorder? Are the findings from a school-based sample generalizable to a clinic-based sample?

If a comprehensive, methodologically sound review of the literature on the efficacy of psychosocial and pharmacological treatments for childhood depression existed, researchers and clinicians would have a better understanding of which treatments are effective in ameliorating depressive symptomatology in children as well as how outcomes covary with study and sample characteristics. Furthermore, a good integrative review would help to guide both intervention efforts and future empirical studies on the treatment of depressive disorders in children.

Treatment Variables

As mentioned previously, there are several variables that must be considered when evaluating whether treatments for depression are effective for children. Central to question of treatment effectiveness is: What treatments are effective for whom and under what conditions? Ultimately, clients are best served when there is a proper fit between the client, the therapist, and the intervention. For example, if an 11-year-old, moderately depressed boy presents for treatment, what is the recommended course of treatment? Based on some of the empirical findings, one might recommend a school-based group intervention such as the Primary and Secondary Control Enhancement training proposed by Weisz et al. (1997). Further, it could be argued that a prerequisite client variable for this type of treatment is an adequate level of intellectual functioning in light of the cognitive underpinnings of Primary and Secondary Control Enhancement training.
However, suppose the child is cognitively impaired, thus precluding him from participating effectively during the treatment group? What if the child were an 8-year-old, severely depressed female youngster? In other words, what can we say about the differential effectiveness of various treatments in light of some of these important variables?

Because a major thrust of this investigation is based upon how treatment study variables moderate outcome, several important variables will be defined and described in the following section, including age, sex, referral source, treatment setting, type of treatment, treatment variables, assessment measures, therapist variables, severity of illness, and comorbid conditions. In addition, available information on how the aforementioned variables impact treatment outcomes will be reviewed.

**Age**

An important variable to consider when selecting the best course of treatment for depression is the client’s age. As mentioned previously, certain therapies might be inappropriate for very young children such as certain cognitive regimens that might presuppose a particular level of cognitive development (e.g., formal operations, Piaget & Inhelder, 1969). No published studies were located that addressed specifically, the potential interaction between age and the treatment of early-onset depression (i.e., adolescent vs. child samples). Yet some interventions target one age group or another (e.g., Coping with Depression Course for Adolescents; Clarke, Lewinsohn, & Hops, 1990). In addition, the empirical findings on the relationship between a child’s age and
general psychotherapy outcome have been equivocal. For example, Weisz et al. (1987) conducted a meta-analysis of 108 treatment studies regarding the overall effectiveness of psychotherapy with children and adolescents. The authors reported that “therapy proved more effective for children (ages 4-12) than for adolescents (ages 13-18)” (p. 542). Weisz et al. reported that the mean effect size for therapy involving children was 0.92 as compared to 0.58 for adolescent clients.

In contrast, Weisz, Weiss, et al. (1995) completed another integrative review (150 different studies) of the effects of psychotherapy with children and adolescents and reported that “treatment outcomes were better for adolescents than for children” (p. 461). For children 11 and younger, the authors reported a mean effect size of 0.48, whereas the mean effect size was 0.65 for adolescents 12 and older. While it is unclear whether the differential inclusion criteria (i.e., “children” defined as 4-12 in Weisz et al., 1987 as compared to 11 and younger in Weisz, Weiss, et al., 1995) impacted the mean effect size calculations, the differential findings leave many questions unanswered. However, Weisz, Weiss, et al. (1995) suggested that the changes in age effect might be attributable to improvements in the effectiveness of adolescent interventions. In any case, the relationship between age and outcome remains unclear, especially with respect to the treatment of child and adolescent depression.

Sex

The interaction between sex and therapy outcome has been investigated by several researchers who have conducted large meta-analytic reviews (e.g., Casey & Berman,
1985; Weisz, Weiss, et al., 1995). Taken together, the findings suggest that psychotherapy is generally more effective for females than males. Weisz et al. (1987) reported a mean effect size of 1.11 for female majority groups and 0.80 for male majority groups. Weisz, Weiss, et al. (1995) reported similar findings, indicating that the mean ES for female samples was 0.71 as compared to 0.43 for male samples. In addition, the authors reported an age/sex interaction wherein the mean ES was 0.86 for adolescent female majority samples as compared to 0.37 for adolescent males. In contrast, the mean effect sizes for the male and female child samples (i.e., 11 and younger) were approximately equal. Thus, while there have been some interesting findings to date, more information is needed to better understand the relationship between sex and depression treatment outcome.

**Referral Source**

How a child or adolescent is referred for treatment is an important variable to consider when evaluating outcome. For example, has the child been clinically referred for treatment due to presenting symptomatology? Or, has the child been recruited by a research team and therefore volunteers to participate as an analog subject? Assuming the treatment is shown to be effective in ameliorating certain symptomatology in volunteer research subjects, the crucial question is whether the “effectiveness” is generalizable to actual clinical cases which may be more difficult to treat successfully. Weisz, Donenberg, Han, and Kauneckis (1995) reported that it is certainly possible that clinic-referred children are more seriously disturbed, that they come from more dysfunctional families, or that they
are more likely to confront stressful life circumstances (e.g., low income levels) than are the often middle class youngsters whose parents give consent for their participation in research studies. (p. 96)

Nevertheless, Weisz, Donenberg, et al. (1995) addressed this question empirically by comparing the mean effect sizes of clinic and analog samples from two large meta-analytic reviews and reported that the differences were not statistically different. Further, based on the results from one of the meta-analyses (i.e., Weisz et al., 1987), the mean effect size of the analog samples (0.76) was smaller than the clinic samples (0.89), indicating that overall, psychotherapy was slightly more effective in treating clinical samples. Thus, the findings to date have been inconsistent regarding the impact of referral source on treatment outcome.

Treatment Setting

The setting in which psychotherapy is conducted is thought to have important implications for therapy outcome (Brent et al., 1997; Weisz, Weiss, & Donenberg, 1992). Treatment is conducted in variety of environments including inpatient units (e.g., hospitals), schools, outpatient clinics, private practice, research laboratories, and correctional settings. Some authors have suggested that treatment delivered during the course of everyday clinical practice as opposed to research or laboratory settings differs in important ways (Seligman, 1995). Clarke (1995) outlined a distinction between "efficacy investigations" and "effectiveness research." According to Clarke, efficacy investigations are controlled laboratory studies, whereas effectiveness research are clinical studies carried out in "broader mental health services systems" (p. 718). While Clarke lamented
how little attention is paid to how the results from controlled clinical trials generalize to applied clinical settings, some investigators have examined this issue.

Weisz et al. (1992) found evidence that controlled research trials tended to produce better treatment outcomes than treatment provided in everyday applied clinical settings. However, the authors cautioned that their findings were based on relatively few clinic studies and that the studies had some methodological problems in comparison to the more rigorous and tightly supervised controlled trials. In addition, Kazdin (1990) suggested that there are major discrepancies between research treatment and clinical practice which make the extrapolation of the findings from one setting to the other problematic. According to the authors, the problem of generalizability may be attributable to factors such as problem foci and the severity of presenting problems. Further, Kazdin (1990) noted that the majority of treatment studies solicit children from school settings and provide treatment in groups, whereas clinically referred children in everyday clinical practice are seen usually on an individual basis.

Furthermore, other investigators have examined whether other variables related to research and clinical settings had differential effects on therapy outcome. For example, Weisz, Donenberg, et al. (1995) questioned whether there is something about clinical settings that undermine the effectiveness of psychotherapy for children (e.g., heavy caseloads, not problem-focused, not as strictly supervised). The authors found evidence that three factors contributed to the “superior” findings found in analog studies including: (a) the prevalent use of behavioral and cognitive-behavioral methods; (b) a reliance on focused, specific therapy techniques; and (c) the provision of structured supervision and
monitoring (e.g., protocol adherence, reliability checks) throughout the trial.

In sum, it appears that certain aspects of the treatment setting have the potential to moderate outcome based on general reviews of the child and adolescent psychotherapy literature. However, it remains to be seen whether some of the aforementioned issues have an impact on interventions for child and adolescent depression.

**Type of Treatment**

A major thrust of this investigation focuses on the relative effectiveness of various treatments for childhood depression. As mentioned previously, several treatments have been developed and implemented for childhood depression including but not limited to psychodynamic, cognitive-behavioral, behavioral, problem-solving, relaxation, social skills training, activity-based, self-control training, interpersonal psychotherapy, family therapy, and pharmacotherapy. However, it appears that no single treatment is clearly superior to the others in ameliorating depressive symptomatology in children. In addition, there are those who suggest that most treatments are virtually equivalent in terms of therapeutic benefit. For example, Rehm (1995) suggested that we abandon the “horse race” comparisons between treatments for depression in adults in light of the fact “that comparison between any two such programs is unlikely to produce important differences in effectiveness” (p. 203). In evaluating the overall effectiveness of psychotherapy, the equivalency of outcome from a variety of treatment methods has been referred to as the “dodo verdict” (Parloff, 1984). While there is some support for the dodo verdict in terms of the efficacy of various forms of psychotherapy for adult
depression and other disorders (e.g., Robinson et al., 1990; Smith, Glass, & Miller, 1980), it remains unclear whether this so-called dodo verdict is applicable to psychotherapy for child and adolescent depression.

For example, three large meta-analytic reviews provided some support for the idea that behavioral treatments are more effective than nonbehavioral methods in treating a variety of child and adolescent psychopathology (Casey & Berman, 1985; Weisz et al., 1987; Weisz, Weiss, et al., 1995). However, the issue is far from settled as some authors have suggested that the differential effects are attributable to artifactual explanations such as methodological differences between studies (e.g., Shirk & Russell, 1992). However, Weiss and Weisz (1995) reported that there was “little support” for such a hypothesis. In any case, there is empirical evidence to justify the comparison of various treatments for child and adolescent problems in light of the aforementioned findings. Further, little is known about the differences between behavioral versus nonbehavioral treatments for childhood depression in particular and how these variables interact with factors (e.g., age-related developmental changes).

Other Treatment Variables

The actual length (i.e., weeks) and frequency (i.e., number of sessions) are important variables to consider when evaluating treatment outcome. An intuitively held assumption in psychotherapy research is that “more therapy is better.” Historically, this assumption has played out in professional practice as the average duration of treatment in clinical settings (27-55 weeks) is substantially longer than it is in research or laboratory
settings (8-10 weeks; cf. Kazdin et al., 1990). However, in light of the present political and financial climate (i.e., cost containment, managed care), the need for empirically validated yet brief interventions for various psychological problems has never been stronger (Weisz et al., 1997). Furthermore, the intuitive assumption that more therapy is better has not been empirically supported (Weisz, Donenberg, et al., 1995). In light of the limited data on this issue, there is a need to better understand the relationship between various treatment variables (e.g., length, frequency) and the effectiveness of treatment for child and adolescent depression (Clarke, 1995).

**Assessment Measures**

As mentioned previously, the process of assessment can have a significant impact on treatment outcome studies because the assessment and diagnostic measures that guide the inclusion and evaluation of subjects are used to determine whether the treatment is efficacious or not. There are a number of methods for assessing depression in youth including: behavioral observations, behavior rating scales (e.g., Child Behavior Checklist CBCL; Achenbach & Edelbrock, 1983), sociometric approaches, diagnostic interviews, and self-report measures, the latter two of which are the most prominent. Several excellent self-report measures have been developed such as: the Children's Depression Inventory (CDI; Kovacs, 1992), the Mood and Feelings Questionnaire (MFQ; Angold, Costello, Pickles, & Winder, 1987), the Reynolds Child Depression Scale (RCDS; Reynolds, 1989), the Center for Epidemiological Studies - Depression Scale (CES-D; Radloff, 1977), the Reynolds Adolescent Depression Scale (RADS; Reynolds, 1987), and
the Beck Depression Inventory (BDI; Beck & Steer, 1987). Some of the more prominent diagnostic interviewing formats include the Schedule for Affective Disorders and Schizophrenia for School-Aged Children (K-SADS; Puig-Antich & Chambers, 1978), the Children’s Depression Rating Scale (CDRS; Poznanski & Mokros, 1996), the Bellevue Index of Depression (BID; Petti, 1978), and the Diagnostic Interview for Children and Adolescents-Revised (DICA-R; Reich & Welner, 1988).

Nonetheless, many investigators (e.g., La Greca, 1990) have stressed the importance of a multimethod protocol in light of the limitations of any one particular measure. At the same time, this process is often more the exception than the rule, and there has been a historical overreliance on a limited number of devices to evaluate depression treatment efficacy. For example, Black-Ceccini (1996) reported that of the 13 treatment studies reviewed, 5 studies relied exclusively on one self-report device to determine efficacy. Further, of the 8 studies that used interview formats, sufficient data to calculate effect sizes were available in only 4 of the studies.

Similarly, there is empirical evidence that depressive symptomatology, as measured by self-report instruments, attenuates over time, and repeated administrations (e.g., Finch, Saylor, Edwards, & McIntosh, 1987; Michael & Merrell, 1998). This attenuation phenomenon has been interpreted in several ways including: (a) an over-endorsement of symptomatology by distressed children upon initial testing (Reynolds, 1986); (b) an expected variation in reported symptomatology due to natural fluctuations in mood over time (Kovacs, 1992); and/or (c) a better understanding of the assessment task during subsequent intervals (Michael & Merrell, 1998).
In sum, it is vital to consider the impact that the measurement technology has upon the determination of treatment efficacy in light of the aforementioned data. While it seems apparent that there is a significant relationship between assessment and treatment outcome, there is a need to examine this relationship in-depth.

**Therapist Variables**

Therapist variables such as level of training, years of experience, and professional discipline have long been considered important when evaluating therapy outcome (Kazdin et al., 1990; Stein & Lambert, 1984). For example, Weisz et al. (1987) reported a significant interaction between level of training and client age, with paraprofessionals and graduate students being more effective with children as opposed to adolescent clients. In contrast, professional therapists were equally effective with both groups (Weisz et al., 1987). In addition, Weisz et al. reported that professional therapists were somewhat more effective than their less trained counterparts (i.e., graduate students, paraprofessionals) with “overcontrolled” problems such as phobias and shyness. In a more recent investigation, Stein and Lambert (1995) conducted a meta-analysis of outcome studies involving within-study comparisons of therapists with different levels of training and experience. The authors concluded that “a variety of outcome sources are associated with modest effect sizes favoring more trained therapists” (p. 182). However, both Weisz et al. and Stein et al. cautioned that it is markedly difficult to come up with an accurate coding scheme to account for all the variables associated with differential training levels across a diverse array of studies (e.g., graduate students being supervised closely by
experienced professionals). Therefore, such methodological problems pose a significant threat to the validity of these findings. In sum, although there is growing evidence that increased levels of training are associated with greater levels of improvement, it is unsettling that “more compelling evidence is not available that demonstrates that graduate training directly relates to enhanced therapy outcomes” (Stein & Lambert, 1995, p. 194). Further, despite some provocative findings, it is unclear whether these trends are evident in the treatment of early-onset depression.

Severity of Illness and Comorbidity

As mentioned previously, childhood depression is neither a discrete nor a homogeneous problem. It varies greatly in both the type and severity of symptom presentation. However, in treatment outcome research, there is usually a specific disorder or symptom that is identified as the dependent variable targeted for change (Hoagwood, Hibbs, Brent, & Jensen, 1995). An important question that needs to be addressed is: Does the type and severity of symptom presentation affect outcome?

Furthermore, in clinical settings, depression is often only one of many clinical syndromes (e.g., anxiety) that may be present in the same individual. A number of potential treatment subjects are often excluded from empirical studies if they present with a variety of comorbid conditions or symptoms even though this is common in everyday clinical settings (Clarke, 1995). These factors can make generalizations from the empirical literature to clinical practice problematic (Hoagwood et al., 1995). Hoagwood et al. (1995) suggested that treatment research needs to adapt to the “real life” situations
that are encountered in clinical practice by considering some of the these important variables such as comorbidity. In general, it would be important to consider whether various treatments for depression are impacted by or have an impact upon, comorbid conditions.

Quality of Study

Various authors have suggested that the quality of a study can impact not only the findings from a particular study, but also the overall findings from an integration or meta-analytic review of several studies (e.g., Wilson & Rachman, 1983; Wortman, 1983). Wilson and Rachman (1983) hypothesized that methodologically weaker studies might skew the results from integrative reviews in the direction of an overestimation of therapy effects. Some aspects of how methodological variables impact ES calculations were examined in the one of the seminal meta-analyses of psychotherapy outcome (Smith et al., 1980). However, Smith et al. did not find compelling evidence that, for example, poor quality studies produced spuriously larger effect sizes in comparison with more methodologically rigorous studies. With regard to child psychotherapy outcome research in particular, Weiss and Weisz (1990) examined the results from large meta-analytic studies to determine whether methodological factors impacted the magnitude of effect sizes. Weiss and Wiesz focused on factors that might compromise internal and external validity such as selection bias, subject attrition, random assignment, measurement flaws, history, therapist variables, and ecological (clinical vs. analog settings) factors (Campbell & Stanley, 1963). Weiss and Wiesz (1990) reported that
together, these factors accounted for two-thirds as much variance as the substantive factors (e.g., type of therapy, age) in the original meta-analysis... [and that] in general, increased experimental rigor was related to larger effect sizes; this argues against the hypothesis that methodologically weak studies have led to an overestimate of therapy effects. (p. 639)

Nonetheless, the quality of each study included in the present analysis will be coded and evaluated to examine whether methodological factors impacted treatment outcome.

Summary

In summary, depression in children and adolescents is a recognized phenomenon that adversely affects the lives of many youth. In light of the prevalence, persistence, and negative outcomes associated with depressive disorders in children and adolescents, several treatments have been developed ranging from psychosocial interventions to pharmacological interventions. Current reviews of the outcome studies for child and adolescent depression have not been adequate for guiding research, policy, and practice because many of the reviews are narrative in nature, methodologically flawed, and/or present vague or conflicting conclusions. Furthermore, while some of the efforts to review and synthesize the body of literature on the treatments for depressed youth have answered some questions, there are many more questions left unanswered. If a comprehensive, methodologically sound review of the literature on the efficacy of psychosocial and pharmacological treatments for childhood depression existed, there would be a better understanding of which treatments are effective in ameliorating depressive symptomatology in children as well as how outcomes covary with study and sample characteristics. Furthermore, a good integrative review would help to guide both

Purpose and Objectives

The purpose of this research project was to determine the overall efficacy of psychosocial and pharmacological treatments for child and adolescent depression. An additional purpose of this inquiry was to examine how certain variables impacted treatment outcomes (e.g., age, sex, type of treatment, quality of study).

The objectives of this study were as follows:

1. To determine the overall efficacy of psychosocial interventions for child and adolescent depression.

2. To determine the overall efficacy of pharmacological treatments for child and adolescent depression.

3. To determine whether there is differential efficacy for psychosocial and pharmacological treatments when the following 11 variables are considered: (a) age, (b) sex, (c) referral source, (d) treatment setting, (e) type of treatment, (f) length and frequency of treatment, (g) assessment measures, (h) therapist variables, (i) severity of illness, (j) comorbidity, and (k) quality of study.

Given the aforementioned purpose and objectives, the following three research questions were addressed in this investigation:

1. What is the overall efficacy of psychosocial interventions for child and adolescent depression?
2. What is the overall efficacy of pharmacological treatments for child and adolescent depression?

3. Is there evidence of differential efficacy for psychosocial and pharmacological treatments when the following 11 variables are considered: (a) age, (b) sex, (c) referral source, (d) treatment setting, (e) type of treatment, (f) length and frequency of treatment, (g) assessment measures, (h) therapist variables, (i) severity of illness, (j) comorbidity, and (k) quality of study?
CHAPTER III

METHODOLOGY

Population and Sample

The population for this investigation included empirical studies on the treatment of child and adolescent depression. The overall sample for this investigation included 37 psychosocial and pharmacological studies targeting child and adolescent depression published between 1980-1998. There were 23 psychosocial studies (36 separate treatments) and 14 pharmacological trials. Case reports and single-subject designs were not included. The articles were located through a comprehensive search strategy including an extensive computer search of databases such as PsycLIT, ERIC, and Medline. Key words and various combinations were used including: depression, depressed, mood, internalizing, child, children, childhood, adolescent, adolescence, juvenile, treatment, therapy, psychotherapy, intervention, pharmacotherapy, pharmacologic, medication, tricyclic, antidepressant, SSRI, TCA, outcome, and meta-analysis. Hand searches of the reference lists from the obtained articles were conducted in an effort to find additional articles and several authors were contacted to inquire about "in press" studies or unpublished manuscripts. Manual searches of a variety of peer-reviewed journals were completed as well, including the Journal of Consulting and Clinical Psychology, the Journal of the American Academy of Child and Adolescent Psychiatry, Behavior Therapy, Archives of General Psychiatry, School Psychology Review, Behavioral Psychotherapy, the Journal of Clinical Child Psychology, the Journal

Finally, a comprehensive search of unpublished theses and dissertations via ProQuest Dissertation Abstracts was completed dating back to 1980. Fifteen potentially relevant dissertations were located, of which 10 were ordered. Several of the dissertations ordered were not received (i.e., four) due to restrictive lending policies at the various institutions. Of the six dissertations received, two were deemed appropriate for coding and included in the final sample of studies.

For a study to be included in the investigation, the effects of a particular treatment on child and adolescent depression had to be examined. The following specific criteria must have been met as well: (a) the study had to be a within- or between-subjects group design; (b) the sample was targeted for intervention based upon “at risk” status (i.e., prevention), presenting depressive symptomatology, or a depressive diagnosis; (c) the subjects targeted for intervention were between the ages of 5-18; (d) the treatment was psychosocial or pharmacological in nature (e.g., group, individual, family, TCA, SSRI); and (e) the study utilized at least one depression outcome measure once the intervention was completed.

Design

An integrative or meta-analytic design was used during the course of this study, whereby the results from related treatment studies can be compared (Glass, 1977). In conducting a meta-analysis, the investigator codes the findings from empirical studies and converts them into an effect size, a common metric upon which comparisons across the
various studies can be made. A commonly used effect size is the standardized mean
difference, that is, the treatment group mean minus control group mean, divided by
control group or pooled standard deviation (Smith et al., 1980).

In terms of what an effect size actually indicates, Cohen (1988) stated that “it is
convenient to use the phrase ‘effect size’ to mean ‘the degree [d] to which the
phenomenon is present in the population,’ or ‘the degree to which the null hypothesis is
false’ ” (pp. 9-10). Effect sizes are typically expressed in standard deviation units (Wolf,
1986). Thus, if a calculated effect size (d) is expressed as .50, this is interpreted as one-
half of a standard deviation difference between the means of the groups under study (e.g.,
treatment versus control groups), indicating that “33% of the combined area covered by
two normal equal-sized equally varying populations is not overlapped” (Cohen, 1988, p.
26). Furthermore, Cohen offered guidelines to interpret the magnitude of an effect size.
According to Cohen (1988), effect sizes between 0.20 and 0.49 are considered “small,”
effect sizes between 0.50 and 0.79 are considered “moderate,” and effect sizes 0.80 or
larger have been described as “large.” Cohen (1988) suggested that these qualitative
interpretations are somewhat arbitrary and should be made with caution, keeping in mind
the particular variables under study. For example, if one is measuring the outcome of a
blood-pressure reduction treatment wherein the standard deviation is three diastolic
points, an effect size of 2.0 would be a large effect, but arguably clinically meaningless
because the effect size is based upon a mean difference of approximately six diastolic
points. As such, the practical implications of such an effect size are marginal at best.
Thus, the scaling of the outcome measures must carefully considered before interpreting
the magnitude of effect sizes

In the present study, a meta-analytic design was used because these designs are often more efficient and effective than narrative reviews in accurately synthesizing large amounts of data from a number of studies (Glass, McGaw, & Smith, 1981). Wolf (1986) suggested that some of the common problems associated with narrative reviews can be addressed by conducting a meta-analysis, including biased selection of studies, differential weighting of studies, misleading interpretations of study findings, and failure to examine moderating variables in the questions being addressed. At the same time, it has been suggested that the standards of methodological rigor in meta-analyses be just as stringent as empirical studies (Wolf, 1986). For example, Glass et al. (1981) grouped criticisms of meta-analyses into four domains including: (a) the diversity across studies (e.g., sample characteristics, measurement technology, treatment variables) makes comparisons inappropriate; (b) the results from poorly designed studies given equal weight with well-designed studies skew the results; (c) studies that produce nonsignificant findings are rarely published, thus creating a biased sample of studies; and (d) studies given too much weight have multiple results.

The aforementioned criticisms of meta-analytic designs can be addressed in various ways. For example, when dealing with the diversity and methodological differences across studies, the researcher is best advised to code certain characteristics from each study (e.g., treatment type, methodology) and examine whether the variables mediate the meta-analytic results (Wolf, 1986). As mentioned previously, analyzing how study characteristics covary with outcomes is a major thrust of the study. Thus, a
systematic coding system was developed and implemented to examine the relationship between the 11 broadly construed variables and the treatment outcomes (i.e., ES).

As described above, Wilson and Rachman (1983) hypothesized that methodologically weaker studies might skew the results from integrative reviews in the direction of an overestimation of therapy effects. Thus, in order to account for the potential problem of giving poorly designed studies equal weight with well-designed studies, the quality of each study was rated on the basis of two criteria: (a) the potential threats to internal validity and (b) the overall validity of the study. These criteria were analyzed to determine whether there were differential outcomes based on study quality.

When dealing with a potentially biased sample of studies, also referred to as the “file drawer” problem (Rosenthal, 1979), it has been suggested that a more comprehensive search strategy be employed along with broader inclusion criteria (Glass et al., 1981). As outlined above, a comprehensive search strategy was employed in the present study, including a review and solicitation of unpublished manuscripts. Although the majority of the studies included in the present analysis were published, a moderate amount of data was included from unpublished manuscripts (e.g., Curry & Wells, 1998) and dissertations (e.g., Hickman, 1994).

With respect to the unequal weighting of studies problem, some authors have suggested averaging multiple results from one study to yield perhaps one or two effect sizes, thus preventing studies with multiple results from having undue influence on the overall estimate of the effect size across studies (Smith et al., 1980). To address this issue in the present investigation, a maximum of two effect sizes per treatment modality
were calculated to prevent unequal weighting of the effects from a particular study with the potential for multiple effect size calculations.

Instrumentation and Analysis

A comprehensive coding sheet was developed so that each of the treatment outcome studies was evaluated based on 11 broadly construed variables including: (a) age, (b) sex, (c) referral source, (d) treatment setting, (e) type of treatment, (f) length and frequency of treatment, (g) assessment measures, (h) therapist variables, (i) severity of illness, (j) comorbidity, and (k) quality of study. The quality of each study was rated based on: (a) potential threats to internal validity and (b) the overall validity of the study. Seven potential threats to internal validity were evaluated on a scale of 0 to 3 (0 = not a plausible threat; 1 = minor threat; 2 = plausible threat; 3 = by itself could explain the findings). The potential threats included: maturation, history, testing, instrumentation, regression, selection bias, and experimental mortality. These threats to internal validity and the associated effects upon treatment outcomes have been described extensively in the literature (e.g., Borg & Gall, 1989; Campbell & Stanley, 1963). The overall validity of study was assessed based on several factors including the aforementioned threats to internal validity, sample size, selection procedures, methodological rigor, and measurement technology. A 5-point Likert scale (5 = excellent; 4 = good; 3 = fair; 2 = inferior; 1 = unacceptable) was used to evaluate each study.

Each study was double-coded by the primary researcher and a second Ph.D. candidate in psychology with extensive experience in coding studies for meta-analytic
designs. The second coder was trained in the specific procedures outlined in this study, and interrater reliability coefficients were calculated for each variable coded. Two formulas were used to calculate the interrater agreement. First, interrater reliability was calculated by dividing the total number of congruent observations (CO) by the total number of observations (TO) and multiplied by 100. Second, Cohen’s Kappa (Cohen, 1960) was calculated for each variable to compensate for the limitations in the first formula since the procedure is used to determine agreement beyond chance levels. After calculating reliability data, disagreements in the coding of certain variables were reconciled through consultation and clarification among the coders. A copy of the coding sheet is included in the Appendix.

The results and variables from the various treatment studies were analyzed by computing a number of effect sizes. Effect sizes were calculated with the assistance of the DSTAT computer software program (Johnson, 1989) whereby the outcome data (i.e., means, standard deviations, sample sizes) were entered. In cases where the means and standard deviations were not provided, effect sizes were computed from other data reported in the studies (e.g., F-ratios or t-statistics). Furthermore, in a large number of pharmacologic studies means, standard deviations, F-ratios, or t-statistics were not reported. Instead, the percentage of improvement in the treatment and placebo-control groups was reported. Treatment response in these studies was determined by whether the subjects met criteria for “improvement” (i.e., change observed on particular measures). However, the criteria for improvement and the measures utilized to determine change varied across the studies. Nonetheless, effect size estimates were calculated by
transforming the difference in proportions (percent improved) between the experimental and placebo-control group subjects. This procedure (probit transformation) has been used in a number of meta-analyses (e.g., Clum, Clum, & Surls, 1993; Miller, 1977) and is described elsewhere (see Smith et al., 1980).

Although this procedure is far from ideal in calculating ES, it is more effective and accurate than simply excluding the study from the analysis. One of the primary limitations of this procedure is the inevitable loss of information that takes place after the outcome data are transformed multiple times into an effect size. Nonetheless, it addresses a criticism (Anderson, 1995) of a previous meta-analysis (Hazell et al., 1995) on the efficacy of using antidepressant medications in children and adolescents whereby effect sizes were calculated from only 6 of the 12 studies due to the limited data contained in the articles. For studies that did not report any data suitable for conversion to effect sizes, but were important to include in the final analysis to increase comprehensiveness (and to reduce selection bias), an effect size estimate of zero was entered. Assigning an effect size of zero is a conservative procedure designed to prevent an inflated overall effect size estimate, when including important, albeit incomplete information from particular studies (Rosenthal, 1984). In the present study, an effect size of zero was entered only when the data could not be converted to an effect size and the authors reported nonsignificant differences between the active and placebo conditions. An effect size of zero was entered for three of the pharmacological studies that reported nonsignificant differences between the active versus the placebo conditions and the data in the articles were insufficient to compute effect sizes.
For between-subjects group designs, the traditional meta-analytic formula, first proposed by Glass (1977), was utilized whereby unweighted effect sizes are calculated by subtracting the control group mean from the treatment group mean, divided by the control group (Glass, 1977) or the pooled standard deviation (standard deviation of treatment and control or Cohen's \( d \); Cohen, 1988). Because the size of the denominator (i.e., standard deviation) in this formula can have a serious impact on the magnitude of effect, a concerted effort was made to obtain the best estimate of the variance of an untreated sample. Thus, in this study, the "pooled" standard deviation included the variance of the treatment group before treatment in addition to the variance of the no-treatment or wait-list control group. This procedure is often helpful in obtaining a more accurate estimate of the variance of an untreated group (Karl White, personal communication, February 19, 1998).

Despite the fact that a large number of meta-analytic reviews have used unweighted effect size calculations (e.g., Casey & Berman, 1985; Smith et al., 1980), there is a recent trend in the literature to use a weighted effect size procedure (e.g., DuPaul & Eckert, 1997; Weisz, Weiss, et al., 1995). The unweighted least squares model (ULS) relies on several statistical assumptions, most notably the assumption of homogeneity of variance. That is, the means and standard deviations (i.e., variance) are equivalent across groups and observations. This assumption is closely related to distribution and sample size factors. In meta-analytic studies, individual observations are the ES. Thus, when aggregating a number of effect sizes (i.e., averaging effect sizes across studies), the assumption of homogeneity of variance is often violated in meta-
analytic reviews because many of the studies often have substantial differences in sample sizes. The weighted least squares (WLS) general linear model has been described as a more reliable way to compute effect size estimates since the procedure corrects for biased estimates derived from smaller samples (Hedges & Olkin, 1985). In brief, this procedure corrects for smaller samples by weighting the effect sizes by the inverse of its variance.

However, the WLS method has also met with criticism because comparisons between previous ULS meta-analytic studies and more recent WLS studies are problematic since it would be difficult to determine whether the differences were related to the data or associated artifactually to the analytical procedures. In the meta-analysis conducted by Weisz, Weiss, et al. (1995), they calculated effect size estimates using both the ULS and WLS methods. Based on the six depression treatment studies included in the analysis, the ULS effect size estimate was 0.67, whereas the WLS effect size estimate was 0.64. Thus, practically speaking, the WLS method had little impact upon the average effect size for depression treatment studies. In light of these data, coupled with the fact that the DSTAT program is based on the ULS model, the unweighted least squares model was utilized in this study to facilitate comparisons across ULS meta-analyses.

For within-subjects group designs, unweighted effect sizes were calculated as well. However, in these studies, the effect size was interpreted as the magnitude of change observed in the experimental group from the pretreatment to posttreatment phases (i.e., intrasubject variance). It was hypothesized that these effect sizes would be larger in general, due to the fact that the magnitude of these effect sizes were not mediated by the measurement of change due to the mere passage of time (i.e., “treatment” effect for
control group subjects). As such, comparisons of within- and between-group effect sizes were inappropriate. However, pre-/post-effect sizes were calculated for the between-group designs to facilitate a useful comparison of treatments across the two designs.

A variety of procedures were used to extract and subsequently analyze the data from early-onset depression treatment studies. Comparisons of the main effects, potential interactions, and other specified variables were conducted. Further, to gain an idea of the robustness of the findings, two additional interpretive aids were used. In addition to Cohen's (1988) guidelines, the magnitude of the overall mean difference effect sizes were interpreted based on 95% confidence intervals (CI) and a Fail Safe N procedure was used. The Fail Safe N procedure, first used by Rosenthal (1979) and described by Cooper (1979), generally denotes the number of studies with null effects that would be needed to reverse the conclusion that a statistically significant relationship exists based on the criterion values, usually $p = .05$ or $.01$ (Wolf, 1986). Orwin (1983) adapted this procedure for meta-analytic designs whereby an effect size criterion value is substituted for the probability values. Orwin suggested that Cohen's (1988) guidelines be used as the criterion values (i.e., effect sizes between 0.20 and 0.49 = small; effect sizes between 0.50 and 0.79 = moderate; effect sizes of 0.80 or greater = large). The Fail Safe N for meta-analytic designs is interpreted as the number of studies with a particular criterion value (e.g., effect sizes = 0.20) that are required to bring down the average effect size to that given criterion value (i.e., small effect). In essence, the Fail Safe N procedure helps to account for one of the criticisms (i.e., selection bias) of meta-analyses in that it provides an index of how many studies with "small" effects (unpublished or not included) that
would be needed to diminish the overall effect size to a "small" effect.

Summary

In summary, a comprehensive sample of studies on the psychosocial and pharmacological treatment of early-onset depression were located through an extensive literature search. Articles that met the inclusionary criteria were subsequently analyzed. The outcome data from 37 outcome studies were extracted and converted into effect sizes. Comparisons of main effects, potential interactions, and other specified variables were conducted. Finally, the data were discussed within the context of a number of interpretive aids often used in meta-analytic designs.
CHAPTER IV

RESULTS

The presentation of results is divided into the following four sections: (a) descriptive statistics, (b) intercoder agreement ratings, (c) main effects, and (d) analysis of study variables.

Descriptive Statistics

Psychosocial Studies

Of the 23 psychosocial studies included in the analysis, 14 (61%) were between-subject group studies with wait-list, placebo, or no-treatment control groups. The remaining nine (39%) psychosocial studies utilized a within-subject (pre/post) group design. The psychosocial studies were published or conducted between 1980 and 1998, with the vast majority of studies published after 1990 (i.e., 87%). There were 36 active treatments across all 23 psychosocial studies and two attention-placebo conditions. The modal treatment regimen was cognitive-behavioral group therapy (n = 11), followed by nondirective, supportive individual therapy (n = 3), social skills group therapy (n = 3), cognitive-behavioral individual therapy (n = 4), and relaxation group therapy (n = 2). The remaining treatments included nondirective supportive group therapy, residential treatment, behavioral group therapy, aerobic exercise, individual relaxation therapy, role playing, family therapy, interpersonal therapy, and various combinations (e.g., group CBT plus parent group or group CBT plus family therapy). There were three prevention
studies, 11 studies in which the subjects were targeted based on depressive symptomatology, and nine studies where the subjects were included based on a depressive diagnosis (i.e., major depressive disorder and/or dysthymic disorder).

In the between-subject group studies, there were 985 subjects between the ages of 7 and 18 years old. The number of subjects in each of the controlled studies ranged from 7 to 152 (median = 57.5). There were five (35%) controlled studies in which the average age of the subjects was above 13 years old, whereas nine (65%) studies included subjects with a mean age of below 13 years old. In the within-subject group designs, there were 391 subjects between the ages of 5 and 18 years old. The number of subjects in each of the pre-/poststudies ranged from 8 to 107 (median = 46). Of the nine pre-/poststudies, six (66%) included subjects with an average age above 13 years old, whereas three (33%) studies included subjects with a mean age below 13 years old. The average percentage of female subjects across all 23 studies was approximately 56%. However, only 3 of the 23 studies (13%) reported separate findings based on sex.

The number of therapy sessions in the controlled trials ranged from 8 to 27 (median = 10.5) and the number of weeks for each treatment ranged from 2 to 12 weeks (median = 8). The number of sessions in the pre-/poststudies ranged from 5 to 36 (median = 11) and the number of weeks for each treatment ranged from 4 to 24 weeks (median = 12). Of the 14 controlled studies, 9 (64%) included follow-up data, ranging from 1 month to 1 year posttreatment (median = 6.5 weeks). Among the nine pre-/posttrials, five (55%) included follow-up data, ranging from 1 month to 1 year posttreatment (median = 4 months). See Table 1 for a descriptive summary.
Table 1

Frequency and Percentage of Study Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of studies reviewed</td>
<td>55</td>
<td>100.0</td>
</tr>
<tr>
<td>Number excluded</td>
<td>18</td>
<td>32.7</td>
</tr>
<tr>
<td>Reasons for exclusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-subject or case report</td>
<td>8</td>
<td>44.5</td>
</tr>
<tr>
<td>Sample not at risk or depressed</td>
<td>3</td>
<td>16.7</td>
</tr>
<tr>
<td>No depression outcome measure</td>
<td>2</td>
<td>11.1</td>
</tr>
<tr>
<td>Insufficient data to compute ES</td>
<td>5</td>
<td>27.7</td>
</tr>
<tr>
<td>Total number of studies included</td>
<td>37</td>
<td>67.3</td>
</tr>
<tr>
<td>Type of study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychosocial treatment</td>
<td>23</td>
<td>62.2</td>
</tr>
<tr>
<td>Between-subject (controlled)</td>
<td>14/23</td>
<td>60.9</td>
</tr>
<tr>
<td>Within-subject (pre/post)</td>
<td>9/23</td>
<td>39.1</td>
</tr>
<tr>
<td>Pharmacological treatment</td>
<td>14</td>
<td>37.8</td>
</tr>
<tr>
<td>Between-subject (controlled)</td>
<td>14/14</td>
<td>100.0</td>
</tr>
<tr>
<td>Within-subject (pre/post)</td>
<td>0/14</td>
<td>0.0</td>
</tr>
<tr>
<td>Year of publication (psychosocial)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980-1984</td>
<td>1</td>
<td>4.3</td>
</tr>
<tr>
<td>1985-1989</td>
<td>2</td>
<td>8.7</td>
</tr>
<tr>
<td>1990-1994</td>
<td>8</td>
<td>34.8</td>
</tr>
<tr>
<td>1995-1998</td>
<td>12</td>
<td>52.2</td>
</tr>
<tr>
<td>Year of publication (pharmacological)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980-1984</td>
<td>3</td>
<td>21.4</td>
</tr>
<tr>
<td>1985-1989</td>
<td>2</td>
<td>14.3</td>
</tr>
<tr>
<td>1990-1994</td>
<td>6</td>
<td>42.9</td>
</tr>
<tr>
<td>1995-1998</td>
<td>3</td>
<td>21.4</td>
</tr>
<tr>
<td>Source of study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Published study</td>
<td>34</td>
<td>91.9</td>
</tr>
<tr>
<td>Dissertation or thesis</td>
<td>2</td>
<td>5.4</td>
</tr>
<tr>
<td>Unpublished manuscript</td>
<td>1</td>
<td>27.7</td>
</tr>
<tr>
<td>Follow-up assessment (psychosocial)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-9 weeks</td>
<td>5/14</td>
<td>35.7</td>
</tr>
<tr>
<td>9-16 weeks</td>
<td>1/14</td>
<td>7.1</td>
</tr>
<tr>
<td>&gt;16 weeks</td>
<td>8/14</td>
<td>57.2</td>
</tr>
</tbody>
</table>
Pharmacological Studies

All 14 pharmacological studies included in the analysis were controlled clinical trials with active medication and placebo conditions, with the exception of one study in which an inactive placebo pill was combined with psychotherapy (Mandoki et al., 1997). All 14 studies included subjects based on a diagnosis of major depressive disorder based on DSM criteria (APA, 1980, 1987, 1994) and information solicited from self-report measures, interviews, and observations. The pharmacological trials were published or conducted between 1981 and 1997 and 64% of studies were published after 1990. There were 441 subjects between the ages of 6 and 19 years old. The number of subjects in each of the pharmacological trials ranged from 6 to 96 (median = 30). There were six (43%) studies in which the subjects were 12 years of age or younger, six (43%) studies in which the subjects were 13 years of age or older, and two (14%) studies in which there were children and adolescents ranging from 7 to 17 years old (Emslie et al., 1997) and 8 to 17 years old (Mandoki et al., 1997). The average percentage of female subjects across all pharmacological studies was approximately 42%. However, none of the studies reported separate findings based on sex.

The types of medications used in the pharmacological trials included imipramine (Tofranil; n = 4), amitriptyline (Elavil; n = 3), desipramine (Norpramin; n = 2), nortriptyline (Pamelor; n = 2), fluoxetine (Prozac; n = 2), and venlafaxine (Effexor; n = 1). The number of weeks for each treatment ranged from 4 to 8 weeks (median = 6). The outcome data in the pharmacological trials were reported at the end of the acute phase of pharmacotherapy (i.e., posttreatment). Follow-up data were not included in the
original articles. See Table 1 for a descriptive summary.

Intercoder Agreement

As described in Chapter III, 25 psychosocial studies were double-coded by the primary researcher and a second Ph.D. candidate in psychology. Although two of the studies rated were not included in the final sample, the intercoder agreement data from these studies were retained to increase the chances of obtaining valid reliability estimates for the entire sample of studies. The pharmacological articles were not double-rated due to the limited number of variables, relative to the psychosocial studies, that were deemed suitable for coding. Two formulas were used to calculate the interrater agreement. First, interrater agreement percentages were calculated by dividing the total number of CO by the total number of observations TO and multiplied times 100. Second, Cohen's Kappa (Cohen, 1960) was calculated for each variable to compensate for the limitations in the first formula. Disagreements in the coding of any variables were reconciled through consultation and clarification among the coders. Interrater agreement rates ranged from 88-100% with an average agreement of 95.3%. Kappa coefficients ranged from 0.76 to 1.00, with an average Kappa reliability coefficient of .91. See Table 2 for a summary of intercoder agreement ratings on key study variables.

Main Effects

Psychosocial Studies

As mentioned previously, one of the primary objects of the current investigation
was to determine the overall efficacy of psychosocial interventions for child and adolescent depression. Tables 3 through 6 provide a summary of key information regarding study main effects. Based on a review of the 14 controlled psychosocial studies (21 treatments) included in the analysis, the overall mean difference effect size at posttreatment was 0.74 (range = 0.03 - 1.84; 95% CI 0.49 - 1.01), indicating that the average child who received treatment for depression was better off than approximately 77% of the children who did not. Of the eight controlled psychosocial studies that reported follow-up data, the mean effect size was 0.64 (range = 0.08 - 1.55; 95% CI 0.32 - 0.95), indicating that the average treated child was better off at follow-up (median = 6.5 weeks) than approximately 74% of control group subjects. Based on the Fail Safe N formula with the criterion level set at 0.20 (i.e., small effect), it would take approximately
Table 3

Overall Mean Effect Sizes at Posttreatment

<table>
<thead>
<tr>
<th>Type of study</th>
<th>N</th>
<th>Mean ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled psychosocial</td>
<td>14 (21 treatments)</td>
<td>.74</td>
<td>0.49 - 1.01</td>
</tr>
<tr>
<td>Intrasubject treatment</td>
<td>(1.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrasubject control</td>
<td>(0.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-/postpsychosocial</td>
<td>9 (15 treatments)</td>
<td>1.14</td>
<td>0.75 - 1.52</td>
</tr>
<tr>
<td>Pharmacological</td>
<td>13</td>
<td>.19</td>
<td>-0.08 - 0.45</td>
</tr>
</tbody>
</table>

Note. Parenthetical ES values reflect amount of change on the dependent measures that was observed in controlled studies using intrasubject (pre/post) variance estimates.

Table 4

Overall Mean Effect Sizes at Follow-up

<table>
<thead>
<tr>
<th>Type of study</th>
<th>N</th>
<th>Mean ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled psychosocial</td>
<td>8 (12 treatments)</td>
<td>.64</td>
<td>0.32 - 0.95</td>
</tr>
<tr>
<td>Pre-/postpsychosocial</td>
<td>5 (8 treatments)</td>
<td>1.26</td>
<td>0.99 - 1.52</td>
</tr>
<tr>
<td>Pharmacological</td>
<td>13</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

56 studies, all with effect sizes of 0.20, to bring down the overall posttreatment mean effect size to 0.20

In regards to the nine pre-/postpsychosocial studies (15 treatments), the overall mean effect size at posttreatment was 1.14 (range = 0.23 - 2.30; 95% CI 0.75 - 1.52). For the five studies that reported follow-up data, the mean effect size at follow-up (median = 36 weeks) was 1.26 (range 0.95 - 1.94; 95% CI 0.99 - 1.52; see Tables 3 and 4). Based on the Fail Safe N formula with the criterion level set at 0.20 (i.e., small effect), it would take approximately 70 studies, all with effect sizes of 0.20, to bring down the overall posttreatment mean effect size to 0.20.
Table 5

Effect Sizes at Posttreatment and Follow-Up for Controlled Psychosocial Studies

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample</th>
<th>Treatment</th>
<th>Post ES</th>
<th>E-ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown et al. (1992)</td>
<td>7 female adolescents (ages 13-16) with depressive symptoms</td>
<td>Aer Exer</td>
<td>0.18</td>
<td>--</td>
</tr>
<tr>
<td>Butler et al. (1980)</td>
<td>56 children (ages 9-12) with depressive symptoms</td>
<td>CBT-G</td>
<td>0.09</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RP-G</td>
<td>0.29</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Attn-Plac</td>
<td>-1.15*</td>
<td>--</td>
</tr>
<tr>
<td>Clarke et al. (1995)</td>
<td>120 adolescents (ages 13-17) with depressive symptoms</td>
<td>CBT-G</td>
<td>0.31</td>
<td>0.10</td>
</tr>
<tr>
<td>Hickman (1994)</td>
<td>9 children (ages 8-11) with depressive symptoms</td>
<td>SS-G</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>Jaycox et al. (1994); Gillham et al. (1995)</td>
<td>143 “at-risk” children (ages 10-13) with depressive symptoms</td>
<td>CBT-G</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>Kahn et al. (1990)</td>
<td>68 children (ages 10-14) with depressive symptoms</td>
<td>CBT-G</td>
<td>1.84</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rix-G</td>
<td>1.30</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBT-I</td>
<td>1.25</td>
<td>0.68</td>
</tr>
<tr>
<td>Lewinsohn et al. (1990)</td>
<td>59 adolescents (ages 14-18) with MDD or Dysthymia</td>
<td>CBT-G</td>
<td>1.11</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBT-G + P</td>
<td>1.45</td>
<td></td>
</tr>
<tr>
<td>Lewinsohn et al. (1996)</td>
<td>96 adolescents (ages 14-18) with MDD or Dysthymia</td>
<td>CBT-G and CBT-G+P</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Liddle &amp; Spence (1990)</td>
<td>31 children (ages 7-11) with depressive symptoms</td>
<td>CBT-G</td>
<td>0.52</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Attn-Plac</td>
<td>-0.51*</td>
<td>-0.33</td>
</tr>
<tr>
<td>Rawson &amp; Tabb (1993)</td>
<td>127 children (ages 8-12) with depressive symptoms</td>
<td>Residential</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Reivich (1996)</td>
<td>152 “at-risk” children (ages 12-14) with depressive symptoms</td>
<td>CBT-G</td>
<td>0.03</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NST-G</td>
<td>0.18</td>
<td>0.08</td>
</tr>
<tr>
<td>Reynolds &amp; Coats (1986)</td>
<td>30 adolescents (ages 14-18) with depressive symptoms</td>
<td>CBT-G</td>
<td>1.48</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rix-G</td>
<td>1.57</td>
<td>1.55</td>
</tr>
<tr>
<td>Stark et al. (1987)</td>
<td>29 children (ages 9-12) with depressive symptoms</td>
<td>CBT-G</td>
<td>0.93</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BPS-G</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Weisz et al. (1997)</td>
<td>48 children (ages 8-12) with depressive symptoms</td>
<td>CBT-G</td>
<td>0.54</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Note. See Table 6 notes for abbreviations.

aAttention-Placebo group effect sizes not included in mean effect size calculations because they were not active treatment conditions.

bFollow-up data reported but not suitable for conversion to effect sizes.
Table 6

Effect Sizes at Posttreatment and Follow-Up for Each Pre- and Postpsychosocial Study

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample Description</th>
<th>Treatment</th>
<th>Post ES</th>
<th>F-ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brent et al. (1997)</td>
<td>107 adolescents (ages 13-18) with MDD</td>
<td>CBT-I</td>
<td>2.30</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SBFT</td>
<td>1.61</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NST-I</td>
<td>2.04</td>
<td>--</td>
</tr>
<tr>
<td>Curry &amp; Wells (1998)</td>
<td>11 adolescents (ages 14-17) with MD or</td>
<td>CBT-G + F</td>
<td>1.41</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Dysthymia and comorbid substance abuse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine et al. (1991)</td>
<td>66 adolescents (ages 13-17) with MDD or</td>
<td>SS-G</td>
<td>0.23</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>Dysthymia</td>
<td>NST-G</td>
<td>0.76</td>
<td>1.16</td>
</tr>
<tr>
<td>King et al. (1990)</td>
<td>46 “at-risk” children (ages 5-11)</td>
<td>SS-G</td>
<td>0.96</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consult</td>
<td>0.25</td>
<td>--</td>
</tr>
<tr>
<td>Mufson et al. (1994); Mufson</td>
<td>14 adolescents (ages 13-16) with MDD</td>
<td>IPT-I</td>
<td>2.02</td>
<td>1.94</td>
</tr>
<tr>
<td>et al. (1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reed (1994)</td>
<td>8 male adolescents (ages 14-19) with MDD</td>
<td>SS-G</td>
<td>1.45</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vostanis et al. (1996)</td>
<td>56 children and adolescents (ages 8-17)</td>
<td>CBT-I</td>
<td>1.30</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>with MDD or Dysthymia</td>
<td>NFI-I</td>
<td>0.89</td>
<td>0.95</td>
</tr>
<tr>
<td>Wagner et al. (1993)</td>
<td>30 female adolescents (ages 9-15) with</td>
<td>NST-I</td>
<td>0.24</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>depressive symptoms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wood et al. (1996)</td>
<td>53 children and adolescents (ages 11-17)</td>
<td>CBT-I</td>
<td>1.33</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>with MDD or DD</td>
<td>Rlx-I</td>
<td>0.30</td>
<td>1.27</td>
</tr>
</tbody>
</table>

Note. Aer Exer=Aerobic Exercise; Alt-Tx=Alternative Treatment; Attn-Plac=Attention-Placebo Group; CBT-G=Cognitive-Behavioral Therapy-Group; CBT-I=Cognitive-Behavioral Therapy-Individual; Consult=Consultation; F=Family Therapy; IPT-I=Interpersonal Psychotherapy-Individual; NFI-I=Non-focused Intervention-Individual; NST-I=Non-Directive Supportive Therapy-Individual; P=Parent; PS-G=Problem-Solving Group; PsyEd-G=Psychoeducational Group; Rlx-I=Individual Relaxation Therapy; Rlx-G=Group Relaxation Therapy; RP-G=Role-Play Group; SS-G=Social Skills Group; SBFT=Structured Behavior Family Therapy.
However, because these effect sizes were calculated based on intrasubject variance, the mean effect size reflects an overall decrease in the amount of depressive sequelae as measured by the various outcome instruments across all studies. Without control group comparisons, it was virtually impossible to estimate how much of the treatment effect, as determined by the dependent measures, was attributable to the treatment as opposed to extraneous factors such as the mere passage of time. Therefore, in order to facilitate a useful comparison of the mean effect sizes of the between- and within-subject group designs, pre-/post-effect sizes were also calculated for the controlled studies. The overall mean effect size for controlled studies using intrasubject variance estimates was 1.23 (range = 0.30 - 2.27; 95% CI .90 - 1.55), which was comparable with the average effect size for the pre-/poststudies (i.e., ES = 1.14). To gather an understanding of the amount of change that occurred without treatment, pre-/post-effect sizes were calculated for control group subjects only. The pre-/post-effect size calculations for control group subjects yielded a mean effect size of 0.37 averaged across the 14 studies. These data approximate the amount of change between the pretreatment and posttreatment phases that could not be explained by or attributed to the therapeutic interventions themselves (see Table 3).

**Pharmacological Studies**

Of the 14 pharmacological trials that compared active versus placebo response rates, effect sizes were entered for all but one study (i.e., Mandoki et al., 1997), due to the fact that the active medication was compared to a placebo plus psychotherapy condition,
thus introducing a confound to the comparison. Moreover, the authors reported that almost 66% of the subjects in both conditions responded positively to the treatment. Nonetheless, the overall mean effect size of the pharmacological studies at posttreatment was 0.19 (range -0.88 - 1.19; 95% CI -0.08 - 0.45), suggesting that the average subject who was administered the active medication moved to the 58th percentile in the distribution of subjects who took the placebo. Follow-up data were not reported in the pharmacological studies (see Tables 3, 4, and 7). Because the magnitude of the mean effect size for pharmacological studies was below Cohen’s suggested guidelines for a “small effect,” the Fail Safe N formula was not calculated.

Analysis of Study Variables

Age

As previously noted, no published studies were located that addressed specifically, the potential interaction between age and the treatment of early-onset depression. To complicate matters, differentiating between “child” and “adolescent” samples can be erroneous and potentially misleading. However, in order to facilitate a useful analysis of the potential interaction of age and treatment outcome, a false dichotomy was created whereby child studies were defined as having a mean age of 12 or younger and adolescent studies were characterized by a sample with a mean age of 13 or older. Although the most appropriate differentiation between child and adolescent samples remains the subject of much debate, the aforementioned definitions appear most frequently in the published literature (e.g., Emslie et al., 1997; Weisz, Weiss, et al., 1995). Further, the
### Table 7

**Effect Sizes Comparing Active Versus Placebo Positive Response Rates for Controlled Pharmacological Studies**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample</th>
<th>Treatment</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boulos et al. (1991)</td>
<td>30 adolescents (ages 15-19) with MDD</td>
<td>Desipramine</td>
<td>0.44</td>
</tr>
<tr>
<td>Emslie et al. (1997)</td>
<td>96 children and adolescents (ages 7-17) with MDD</td>
<td>Fluoxetine</td>
<td>0.59</td>
</tr>
<tr>
<td>Geller et al. (1990)</td>
<td>31 adolescents (ages 12-17) with MDD</td>
<td>Nortriptyline</td>
<td>-0.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SMDES-CDRS</td>
<td>(0.49)</td>
</tr>
<tr>
<td>Geller et al. (1992)</td>
<td>50 children (ages 6-12) with MDD</td>
<td>Nortriptyline</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SMDES-CDRS</td>
<td>(1.12)</td>
</tr>
<tr>
<td>Hughes et al. (1990)</td>
<td>27 children (ages 6-12) with MDD and Comorbid Anxiety or CD/ODD</td>
<td>Imipramine</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDD + Anx</td>
<td>(1.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDD + CDD/ODD</td>
<td>(-0.88)</td>
</tr>
<tr>
<td>Kashani et al. (1984)</td>
<td>9 children (ages 9-12) with MDD</td>
<td>Amitriptyline</td>
<td>0.00</td>
</tr>
<tr>
<td>Kramer &amp; Feiguine (1981)</td>
<td>20 adolescents (ages 13-17) with MDD</td>
<td>Amitriptyline</td>
<td>0.00</td>
</tr>
<tr>
<td>Kutcher et al. (1994)</td>
<td>60 adolescents (ages 15-19) with MDD</td>
<td>Desipramine</td>
<td>0.34</td>
</tr>
<tr>
<td>Kye et al. (1996)</td>
<td>22 adolescents (ages 12-17) with MDD</td>
<td>Amitriptyline</td>
<td>1.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SMDES-HAM-D</td>
<td>(0.53)</td>
</tr>
<tr>
<td>Mandoki et al. (1997)</td>
<td>33 children and adolescents (ages 8-17) with MDD</td>
<td>Venlafaxine</td>
<td>--</td>
</tr>
<tr>
<td>Petti &amp; Law (1982)</td>
<td>6 children (ages 6-12) with MDD</td>
<td>Imipramine</td>
<td>0.12</td>
</tr>
<tr>
<td>Preskorn et al. (1987)</td>
<td>22 children (ages 6-12) with MDD</td>
<td>Imipramine</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% change CDRS</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Puig-Antich et al. (1987)</td>
<td>38 children (ages 6-12) with MDD</td>
<td>Imipramine</td>
<td>-0.31</td>
</tr>
<tr>
<td>Simeon et al. (1990)</td>
<td>30 adolescents (ages 13-18) with MDD</td>
<td>Fluoxetine</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Note.** Anx=Anxiety; CDE=Children’s Depression Inventory; CDRS=Child Depression Rating Scale; HAM-D=Hamilton Rating Scale for Depression; MDD=Major Depressive Disorder; ODD=Oppositional Defiant Disorder; SMDES=Standard Mean Difference Effect Size.

*Zero entered due to insufficient data to compute ES estimates and nonsignificant differences reported for active versus placebo comparisons. Parenthetical ES values reflect separate calculations based on data from dependent measures contained in the articles or outcomes based on interactional variables.
age-by-treatment outcome interactions were discussed within the context of each type of study. That is, separate findings were reported for controlled psychosocial, pre-/postpsychosocial, and pharmacological studies.

In the nine controlled psychosocial studies where the average age of the subjects was 12 or younger, the mean effect size was 0.65 (95% CI 0.34 - 0.95). In contrast, the five controlled adolescent studies yielded an average effect size of 0.93 (95% CI 0.36 - 1.49). In regards to pre-/postpsychosocial studies, the mean effect size for child samples was 0.73 (95% CI 0.14 - 1.30), whereas the mean effect size for the six adolescent studies was 1.35 (95% CI 0.83 - 1.85).

Regarding to the pharmacological trials, there were six studies in which the average age was 12 or younger, seven studies in which the average age was 13 or older, and one study in which the ages ranged from 7 to 17 years (Emslie et al., 1997). However, the authors stratified the sample by age (≤ 12 and ≥ 13) and reported no differences in outcome based on age. Therefore, age-by-treatment outcome data from this study were included in the child and adolescent analyses. For child pharmacotherapy trials, the mean effect size was 0.15 (95% CI -0.12 - 0.42), whereas adolescent pharmacological studies yielded an average effect size of 0.28 (95% CI -0.24 - 0.79). See Table 8 for a summary of data on age and mean effect size.

**Sex**

As described above, the interaction between sex and therapy outcome has been investigated. However, despite some interesting trends, the relationship between sex and
Table 8

Mean Effect Sizes Based on Average Age of Sample

<table>
<thead>
<tr>
<th>Type of study</th>
<th>N</th>
<th>Mean ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled psychosocial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child studies</td>
<td>9</td>
<td>0.65</td>
<td>0.34 - 0.95</td>
</tr>
<tr>
<td>Adolescent studies</td>
<td>5</td>
<td>0.93</td>
<td>0.36 - 1.49</td>
</tr>
<tr>
<td>Pre-/postpsychosocial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child studies</td>
<td>3</td>
<td>0.73</td>
<td>0.14 - 1.30</td>
</tr>
<tr>
<td>Adolescent studies</td>
<td>6</td>
<td>1.35</td>
<td>0.83 - 1.85</td>
</tr>
<tr>
<td>Pharmacological</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child studies</td>
<td>7</td>
<td>0.15</td>
<td>-0.12 - 0.42</td>
</tr>
<tr>
<td>Adolescent studies</td>
<td>(8)</td>
<td>0.28</td>
<td>-0.24 - 0.79</td>
</tr>
</tbody>
</table>

Note. Parenthetical N refers to the data from the Emslie at al. (1997) study which was included in both classifications.

outcome is unclear, especially in relation to the treatment of early-onset depression.

Nonetheless, given the epidemiological data which indicate that the female-to-male ratios for depressive illness begin to approximate the base rates for adult depressive disorders (i.e., approximately 2:1) during middle to late adolescence, it would be important to know whether there are trends of differential efficacy based on sex.

Exploratory correlational analyses revealed little in terms of an interaction between sex and psychosocial depression treatment outcome. For example, the Pearson product-moment correlation between the percentage of female subjects and the mean effect size in controlled studies was 0.16 \((p = .48)\) and explained approximately 2.5\% of the variance. Furthermore, the correlational data from pharmacological trials were even more dismal, which could be attributed, in part, to the truncated range of female percentages observed across medication studies.

However, when the data from psychosocial studies were analyzed by calculating
effect sizes using different percentages of female subjects as criteria, there were some noteworthy findings. For example, when the percentage of female subjects in controlled studies was 60 or greater (i.e., the majority), the mean effect size was 0.90, as compared to an effect size of 0.63 when the percentage of female subjects was below 60. Similarly, in pre-/poststudies, when the percentage of female subjects was 60 or greater, the average effect size was 1.20, whereas the mean effect size was 1.04 when the percentage of female subjects dropped below 60. See Table 9 for a summary of data on sex and effect size.

Referral Source

As mentioned previously, the interaction between referral source and treatment efficacy has important implications for clinicians and researchers. Based on the 36 psychosocial treatment regimens (23 studies) included in this review, there were 15 clinically referred and 21 nonreferred groups. Most of the clinically referred samples occurred in the context of within-subject designs (i.e., 12), whereas there was a predominance of nonreferred samples in the controlled trials (i.e., 18). In regard to pharmacological interventions, all of the samples were clinically referred children with a diagnosis of major depressive disorder. Thus, the interaction between referral source and pharmacotherapy was not examined.

The mean effect size for clinically referred subjects in controlled trials was 0.44, whereas the average effect size for nonreferred youngsters in controlled trials was 0.79. Comparatively, the mean effect size was 1.20 for clinically referred children in pre-/poststudies. In the case of nonreferred children and adolescents in within-subject designs,
Table 9

Interaction Between Effect and Sex in Psychosocial Studies

<table>
<thead>
<tr>
<th>Type of study</th>
<th>N</th>
<th>Mean ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 60% female</td>
<td>8</td>
<td>0.90</td>
</tr>
<tr>
<td>&lt; 60% female</td>
<td>6</td>
<td>0.63</td>
</tr>
<tr>
<td>Pre-/poststudies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 60% female</td>
<td>5</td>
<td>1.20</td>
</tr>
<tr>
<td>&lt; 60% female</td>
<td>4</td>
<td>1.04</td>
</tr>
</tbody>
</table>

The average effect size was 0.89. See Table 10 for a summary of findings.

Treatment Setting

In the present investigation, 36 psychosocial treatments were delivered in school settings, outpatient clinics, and inpatient hospitals. The school (n = 15; 71%) was the modal setting in controlled psychosocial trials, whereas the outpatient clinic (n = 12; 80%) was the predominant setting in pre-/poststudies. In terms of pharmacological studies, there were six inpatient trials, seven outpatient trials, and one study in which the treatment setting could not be accurately determined but included a number of outpatients (i.e., Simeon et al., 1990).

The average effect size for controlled studies conducted in schools was 0.75. Controlled psychosocial treatments conducted in outpatient settings (n = 4) yielded a mean ES of 0.92. In terms of the pre-/poststudies, outpatient treatments produced an average ES of 1.22, whereas the two school-based interventions yielded a mean ES of 0.61. Medication trials conducted in outpatient settings (effect sizes = 0.26) produced a
Table 10

Interaction Between Effect Size and Referral Source in Psychosocial Studies

<table>
<thead>
<tr>
<th>Type of study</th>
<th>N (treatments)</th>
<th>Mean ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled studies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Referred</td>
<td>3</td>
<td>0.44</td>
<td>-0.25 - 1.12</td>
</tr>
<tr>
<td>Nonreferred</td>
<td>18</td>
<td>0.79</td>
<td>0.50 - 1.08</td>
</tr>
<tr>
<td>Pre-/poststudies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Referred</td>
<td>12</td>
<td>1.20</td>
<td>0.74 - 1.66</td>
</tr>
<tr>
<td>Nonreferred</td>
<td>3</td>
<td>0.89</td>
<td>-0.61 - 2.38</td>
</tr>
</tbody>
</table>

higher mean effect size than inpatient regimens (effect sizes = -0.06). See Table 11 for a summary of data on treatment setting and effect sizes.

Type of Treatment

As reported above, one of the major objectives of this investigation was to examine the differential effectiveness of a variety of treatment modalities in ameliorating depressive sequelae in children and adolescents. As presented in Table 12, the modal psychosocial intervention for depressed youth was cognitive-behavioral group therapy (n = 11; 31%), which yielded an average effect size of 0.69. Other common interventions included nondirective supportive individual therapy (n = 3; 8%; Mean effect size = 1.05), cognitive-behavioral individual therapy (n = 3; 8%; Mean effect size = 1.64), social skills group therapy (n = 3; 8%; Mean effect size = 0.88), and relaxation group therapy (n = 2; 6%; Mean effect size = 1.43). Further, there were two attention-placebo conditions that yielded an average effect size of -0.83. See Tables 12 and 13 for a summary of data on treatment type and effect size.
Table 11

Mean Effect (ES) Based on Treatment Setting

<table>
<thead>
<tr>
<th>Type of study</th>
<th>N (treatments)</th>
<th>Mean ES</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled psychosocial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>15</td>
<td>0.75</td>
<td>0.42 - 1.08</td>
</tr>
<tr>
<td>Outpatient</td>
<td>4</td>
<td>0.92</td>
<td>0.18 - 1.62</td>
</tr>
<tr>
<td>Inpatient</td>
<td>2</td>
<td>0.29</td>
<td>-1.10 - 1.68</td>
</tr>
<tr>
<td>Pre-/postpsychosocial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>2</td>
<td>0.61</td>
<td>-3.90 - 5.11</td>
</tr>
<tr>
<td>Outpatient</td>
<td>13</td>
<td>1.22</td>
<td>0.80 - 1.64</td>
</tr>
<tr>
<td>Pharmacological</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatient</td>
<td>8</td>
<td>0.26</td>
<td>-0.20 - 0.73</td>
</tr>
<tr>
<td>Inpatient</td>
<td>6</td>
<td>-0.06</td>
<td>-0.13 - 0.12</td>
</tr>
</tbody>
</table>

Table 12

Mean Effect Size by Treatment Modality in Controlled Psychosocial Studies

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N (treatments)</th>
<th>Mean ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerobic exercise</td>
<td>1</td>
<td>0.18</td>
</tr>
<tr>
<td>Attention placebo</td>
<td>2</td>
<td>-0.83</td>
</tr>
<tr>
<td>Behavioral Problem-solving</td>
<td>1</td>
<td>0.68</td>
</tr>
<tr>
<td>CBT group</td>
<td>11</td>
<td>0.69</td>
</tr>
<tr>
<td>CBT group + family</td>
<td>1</td>
<td>1.45</td>
</tr>
<tr>
<td>CBT-individual</td>
<td>1</td>
<td>1.25</td>
</tr>
<tr>
<td>Nondirective support group</td>
<td>1</td>
<td>0.18</td>
</tr>
<tr>
<td>Residential</td>
<td>1</td>
<td>0.40</td>
</tr>
<tr>
<td>Relaxation group</td>
<td>2</td>
<td>1.43</td>
</tr>
<tr>
<td>Role play</td>
<td>1</td>
<td>0.29</td>
</tr>
<tr>
<td>Social skills group</td>
<td>1</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Note. CBT=Cognitive-Behavioral Therapy.
Table 13

Mean Effect Size by Treatment Modality in Pre- and Poststudies

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N (treatments)</th>
<th>Mean ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBT group + family</td>
<td>1</td>
<td>1.41</td>
</tr>
<tr>
<td>CBT-individual</td>
<td>3</td>
<td>1.64</td>
</tr>
<tr>
<td>SB family therapy</td>
<td>1</td>
<td>1.61</td>
</tr>
<tr>
<td>Individual consultation</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>Interpersonal therapy-ind</td>
<td>1</td>
<td>2.02</td>
</tr>
<tr>
<td>Nondirective individual</td>
<td>3</td>
<td>1.05</td>
</tr>
<tr>
<td>Relaxation-individual</td>
<td>1</td>
<td>0.30</td>
</tr>
<tr>
<td>Role play</td>
<td>1</td>
<td>0.29</td>
</tr>
<tr>
<td>Social skills group</td>
<td>3</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Note. CBT=Cognitive-Behavioral Therapy; Ind=Individual; SB Family Therapy=Structured Behavioral Family Therapy.

In regards to the types of medications used in the pharmacological trials, the TCA imipramine was used most often (n = 4; 29%) and yielded an average effect size of 0.02. Other TCAs administered in medication trials included amitriptyline (n = 3; 21%; Mean effect size = 0.40), desipramine (n = 2; 21%; Mean effect size = 0.39), and nortriptyline (n = 2; 21%; Mean effect size = -0.15). The SSRI fluoxetine was used twice and produced an average effect size of 0.30. See Table 14 for a summary of medication type and effect size.

Other Treatment Variables

Although there is a need to better understand the relationship between the length and frequency of treatment and the efficacy of child and adolescent depression interventions, exploratory analyses did not yield many compelling trends. For example,
Table 14

Mean Effect Size by Medication in Pharmacological Studies

<table>
<thead>
<tr>
<th>Medication</th>
<th>N (treatments)</th>
<th>Mean ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amitriptyline (Elavil)</td>
<td>3</td>
<td>0.40</td>
</tr>
<tr>
<td>Desipramine (Norpramin)</td>
<td>2</td>
<td>0.39</td>
</tr>
<tr>
<td>Fluoxetine (Prozac)</td>
<td>2</td>
<td>0.30</td>
</tr>
<tr>
<td>Imipramine (Tofranil)</td>
<td>4</td>
<td>0.02</td>
</tr>
<tr>
<td>Nortriptyline (Pamelor)</td>
<td>2</td>
<td>-0.15</td>
</tr>
</tbody>
</table>

the Pearson product-moment correlation between the standard mean difference effect size and the number of therapy sessions was -0.03 ($p = .86$). However, there was a modest negative relationship between the number of weeks in therapy and the mean effect size ($r = -0.40; p = .07$). That is, as the number of treatment weeks increased, the average effect size decreased. Moreover, in terms of the pharmacological interventions, the problem of truncated range (i.e., 4-8 weeks) precluded a meaningful analysis between the duration of treatment and outcome.

**Assessment Measures**

As described previously, the process of assessment has a significant impact on treatment outcome studies because the assessment and diagnostic measures that guide the inclusion and evaluation of subjects are used to determine the effectiveness of the interventions. Of the 23 psychosocial studies, a single self-report measure was used to evaluate the treatment in 8 (35%) of them. Further, two other investigations used a combination of two self-report measures to assess outcome. The most commonly used
In terms of combining assessment procedures, 9 of the 23 studies (39%) used both self-report measures and diagnostic interview schedules. The modal interview schedule used was the K-SADS ($n = 7$) followed by the CDRS ($n = 3$) and BID ($n = 3$). Measures of anxiety (e.g., Revised Children’s Manifest Anxiety Scale; Reynolds & Richmond, 1978) were used in three of the studies and the CBCL was used in two trials. There was also the sporadic use of other instruments that purport to measure constructs such as self-esteem, global functioning, and cognitive distortions (Butler et al., 1980).

Although the interaction between assessment technology and outcome was identified as an important issue to examine, it proved difficult to analyze logistically. However, in a number of studies, depression outcome measures were used in concert. Thus, practical comparisons of two different measures within the same study revealed some interesting findings. As shown in Table 15, the assessment strategy often had a noteworthy impact on the ES estimates. For example, Kahn et al. (1990) used the RADS and the CDI in their study to evaluate the efficacy of cognitive-behavioral group therapy. The mean effect size based on the CDI was 1.52, whereas the mean effect size based on the RADS was higher at 2.17. In another study, Reynolds et al. (1986) used the BDI and the RADS to assess the effectiveness of cognitive-behavioral group therapy. Based on the data, the findings were much more convergent, with the mean effect size on the BDI at 1.50 and a mean effect size of 1.46 on the RADS.
Table 15

Within-Study Effect Size Comparisons of Assessment Measures

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment</th>
<th>Measure</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarke et al. (1995)</td>
<td>CBT-G</td>
<td>CES-D</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HRSD</td>
<td>0.28</td>
</tr>
<tr>
<td>Jaycox et al. (1994)</td>
<td>CBT-G</td>
<td>CDI</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RCDS</td>
<td>0.35</td>
</tr>
<tr>
<td>Kahn et al. (1990)</td>
<td>CBT-G</td>
<td>CDI</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RADS</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>Rlx-G</td>
<td>CDI</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>RADS</td>
<td></td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>CBT-I</td>
<td>CDI</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>RADS</td>
<td></td>
<td>1.47</td>
</tr>
<tr>
<td>Lewinsohn et al. (1990)</td>
<td>CBT-G</td>
<td>CES-D</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BDI</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CES-D</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BDI</td>
<td>1.29</td>
</tr>
<tr>
<td>Reynolds et al. (1986)</td>
<td>Rlx-G</td>
<td>BDI</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RADS</td>
<td>1.56</td>
</tr>
<tr>
<td>Preskorn et al. (1987)</td>
<td>Imipramine</td>
<td>CDI</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CDRS</td>
<td>0.06</td>
</tr>
<tr>
<td>Weisz et al. (1997)</td>
<td>CBT-G</td>
<td>CDI</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CDRS-R</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Note. BDI=Beck Depression Inventory; CBT-G(I)=Cognitive-Behavioral Therapy-Group (Individual; F=Family); CDI=Children’s Depression Inventory; CDRS Children’s Depression Rating Scale (R=Revised); CES-D=Center for Epidemiological Studies--Depression Scale; HRSD=Hamilton Rating Scale for Depression; RADS=Reynolds Adolescent Depression Scale; RCDS=Reynolds Child Depression Scale; Rlx-G=Relaxation Group.
Therapist Variables

As previously described, therapist variables such as level of training, years of experience, and professional discipline are important to consider when evaluating therapy outcome. However, due to limited information contained in the articles, only the therapist's level of training for psychosocial interventions was coded that included (a) paraprofessional therapists (e.g., teachers), (b) graduate trainees, and (c) professional therapists.

Of the 21 controlled psychosocial treatments, graduate students were most often in charge of the therapeutic interventions \( (n = 15; 71\%) \) and produced a mean effect size of 0.69. Professional therapists led four of the controlled treatments (19%), yielding an average effect size of 1.18. There were only two controlled studies in which paraprofessional therapists led the treatment, producing a mean effect size of 0.29.

In terms of the 15 pre-/posttreatments, professional therapists led 11 of them (73%) and produced a mean effect size of 1.29. In the remaining four pre-/posttreatments (27%), graduate students and paraprofessional therapists yielded mean effect sizes of 0.85 and 0.61, respectively. See Table 16 for a summary of data on therapist variables and effect size.

Severity of Illness and Comorbidity

As mentioned above, it is important to know whether treatment efficacy is affected by the severity of the depressive illness. Thus, three levels of severity were coded and analyzed including: (a) "at risk" status (i.e., prevention), (b) presenting
Table 16

Mean Effect Size by Level of Training in Psychosocial Studies

<table>
<thead>
<tr>
<th>Type of study</th>
<th>Training level</th>
<th>N</th>
<th>Mean ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled psychosocial</td>
<td>Professional</td>
<td>4</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>Graduate student</td>
<td>15</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Paraprofessional</td>
<td>2</td>
<td>0.29</td>
</tr>
<tr>
<td>Pre-/postpsychosocial</td>
<td>Professional</td>
<td>11</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>Graduate student</td>
<td>2</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Paraprofessional</td>
<td>2</td>
<td>0.69</td>
</tr>
</tbody>
</table>

depressive symptomatology, and (c) depressive diagnosis (major depressive disorder).

While there were different levels of severity in the psychosocial studies, the pharmacological studies uniformly assessed the subjects as suffering from major depressive disorder.

Of the 21 controlled psychosocial treatments, three were determined to be prevention regimens with a mean effect size of 0.17. There were 15 treatments based on depressive symptomatology that yielded an average effect size of 0.81, whereas the remaining 3 treatments purported to address a diagnosis of major depressive disorder and produced a mean effect size of 0.98.

In terms of the 15 pre-/postpsychosocial interventions, 12 of the treatments focused on depressive diagnoses, 2 were designed to address at-risk status, and 1 intervention targeted depressive symptomatology. The pre-/posttreatments for major depressive disorder produced a mean effect size of 1.32, the prevention treatments yielded
an average effect size of .49, and the single symptomatology-based intervention yielded an average effect size of 0.24. Refer to Table 17 for a summary of data on level of severity and effect size.

Unfortunately, the data regarding comorbidity in the studies were limited and did not facilitate many useful analyses. Although descriptive data regarding comorbid conditions were provided in a few studies, the data were not suitable for coding to investigate potential interactions. Further, in the single psychosocial study that specifically addressed comorbid conditions (i.e., Curry & Wells, 1998; depression and substance abuse), substance abuse data were not available at the time of this writing. However, one finding is important to note. In a pharmacological trial in which the efficacy of imipramine was examined (Hughes et al., 1990), the authors reported an interaction between major depressive disorder, comorbid conditions, and outcome. Patients with major depressive disorder and comorbid anxiety disorders responded much better to the medication (Mean effect size = 1.02) than those with major depressive disorder and comorbid conduct disorder and/or oppositional defiant disorder (Mean effect size = -0.88). Thus, while there appeared to be null main effects, when the interaction between comorbid conditions was taken into account, the data were much more practical and useful.

Quality of Study

As described above, several researchers have suggested that there is a relationship between methodological factors (i.e., study quality) and effect size estimates. The quality
Table 17

Mean Effect Size by Level of Severity in Psychosocial Studies

<table>
<thead>
<tr>
<th>Type of study</th>
<th>Level of severity</th>
<th>N (treatments)</th>
<th>Mean ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled psychosocial</td>
<td>Prevention</td>
<td>3</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>Symptomatology</td>
<td>15</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Diagnosis</td>
<td>3</td>
<td>0.98</td>
</tr>
<tr>
<td>Pre-/postpsychosocial</td>
<td>Prevention</td>
<td>2</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Symptomatology</td>
<td>1</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>Diagnosis</td>
<td>12</td>
<td>1.32</td>
</tr>
</tbody>
</table>

of each study was rated based on: (a) potential threats to internal validity and (b) the overall validity of the study. Evaluative guidelines for study quality were based on a 5-point Likert scale (1 = unacceptable, 2 = inferior, 3 = fair, 4 = good, 5 = excellent).

The standard mean difference effect size estimates were cross-tabulated with the overall validity ratings for each study. Based on a linear regression model, the correlation between study quality and effect size was 0.67 ($p < .0001$). That is, as the study quality increased, so did the standard mean difference effect size. See Figure 2 for a graphic representation of these data.
Figure 2. Relationship between study quality and standard mean difference effect sizes.
CHAPTER V
DISCUSSION

The three research questions that were posed in this investigation addressed: (a) the overall efficacy of psychosocial interventions for child and adolescent depression, (b) the overall efficacy of pharmacological treatments for child and adolescent depression, and (c) whether there was evidence of differential efficacy for psychosocial and pharmacological treatments when a number of variables were considered. This chapter includes a detailed interpretation of the results followed by a review of the clinical implications of the findings. The limitations of this investigation are discussed and suggestions for future research are presented.

Main Effects

Psychosocial Studies

The results of this comprehensive review indicate that overall, psychosocial treatments for early-onset depression lead to a substantial reduction in depressive sequelae in children and adolescents regardless of whether the experimental design was a between- or within-group study. The overall mean difference effect size for controlled studies was 0.74 at posttreatment, indicating that the average child who received treatment for depression was better off than approximately 77% of the children who did not. Furthermore, the treatment effects were durable over a brief period of time.
(median = 6.5 weeks), as the mean effect size was 0.64 for the controlled studies that reported follow-up data.

A practical interpretation of the overall effect sizes for controlled studies suggests that the effects were robust and clinically meaningful. That is, a mean change of 0.74 standard deviation units across a variety of instruments that purport to measure depressive sequelae would move a number of youth in the treated group from a clinical range to a nonclinical range. For example, Stark et al. (1987) reported that the mean CDI score of the treated group was 8.04 at posttreatment, as compared to 18.60 for the control group. Thus, with a standard deviation of approximately 10 on the CDI, an effect size of 0.74 would account for almost an 8-point difference on the CDI scores at posttreatment. This difference, in turn, moved a large group of treated subjects into the nonclinical range (i.e., < 11) based on the suggested interpretation of CDI scores by the instrument’s author (Kovacs, 1992).

These findings mirror global reviews of psychotherapy with children whereby the treated youngsters fared much better than wait-list or no-treatment control subjects at posttreatment and follow-up (effect size = 0.71; Casey & Berman, 1985; effect size = 0.88; Kazdin et al., 1990; effect size = 0.79; Weisz et al., 1987; effect size = 0.71; Weisz, Weiss, et al., 1995). These data are also convergent with the findings from more circumscribed meta-analytic reviews that support the efficacy of social skills interventions for depressed youth (effect size = 0.76; Black-Cecchini, 1996), cognitive-behavioral therapy for adolescents (effect size = 1.02; Reinecke et al., 1998), and a smaller sample of depression treatments (effect size = 0.67; Weisz, Weiss et al., 1995).
Furthermore, these findings are consistent with the data from a large number of studies that support the efficacy of psychosocial treatments for depressed adults.

In regards to the pre-/postpsychosocial studies, the overall mean effect size at posttreatment was 1.14, which indicates that the treated subjects experienced significant reductions in depression at posttreatment. These findings were also durable over time (median = 36 weeks), with a mean effect size of 1.26 for the studies in which follow-up data were reported. However, as mentioned previously, these effect sizes were calculated based on intrasubject variance and comparisons between the overall effect size of the controlled and pre-/poststudies were inappropriate. Therefore, in order to facilitate a useful comparison of the mean effect size of the between- and within-subject group designs, pre-/post effect sizes were also calculated for the controlled studies. The overall mean ES for controlled studies using intrasubject variance estimates was 1.23, which is comparable to the mean effect size for pre-/poststudies (1.14).

Given these roughly commensurate values, a case could be made that the treatments in pre-/poststudies led to equally substantial reductions in depressive sequelae despite the methodological limitation of not having wait-list or no-treatment control group comparisons. Further, if the data regarding the average reduction of depression in control-group subjects (0.37) is subtracted from the pre-/post-effect size estimates of treated subjects in between-group studies (1.23), the resulting value (0.86) approximates the overall mean effect size for controlled studies (0.74). Therefore, although the overall effect size for pre-/poststudies was relatively inflated because there were no wait-list or control group comparisons, the similar pre-/postvalues across methodological designs
augment the overall findings that psychosocial interventions are efficacious in the
treatment of early-onset depression.

Pharmacological Studies

The results of this meta-analytic review indicate that in general, pharmacological
treatments for early-onset depression do not lead to a substantial reduction in depressive
sequelae in children and adolescents. The overall mean effect size at posttreatment was
0.19, suggesting that the average subject who was administered the active medication
moved to the 58th percentile in the distribution of subjects who took a pharmacologically
inert placebo. It is unclear whether these effects were maintained over time as follow-up
data were not reported in the pharmacological studies. Further, the average placebo
response rate across all of the studies was high (43%; range 17% - 68%), indicating that
much of the derived benefits from pharmacotherapy were produced by the children’s
expectations of the treatment as opposed to the active medication itself.

In comparison to the only other meta-analytic review located on the efficacy
pharmacotherapy for depressed youth (ES = 0.35; Hazell et al., 1995), these findings were
somewhat more discouraging but roughly consistent, suggesting that antidepressant
medications are not substantially superior to placebo in treating depressed youth. At the
same time, these data are inconsistent with the empirical evidence that the use of
psychotropic medications in adults suffering from depression is efficacious, with an
average positive response rate between 65% and 75% (Elkin, 1994; Speier et al., 1995;
Thase & Kupfer, 1996).
However, a number of important points must be made about the interpretation of these main effects. First, as Conners (1992) aptly pointed out, "various methodological problems limit the conclusions" of pharmacological studies since there are substantial differences in the criteria for improvement, instrumentation, and selection procedures. Thus, in the present investigation, the pharmacological findings represent a synthesis of data that are, in many ways, flawed (i.e., difficulty transforming response rates to effect sizes). For example, in one study improvement might be defined as evidencing a 50% reduction in depressive symptoms as measured by the CDI, whereas one of the criteria for improvement in another study might be a score of 26 or less on the CDRS. The resulting dichotomous determinations of "improved" versus "not improved" in the active and placebo conditions are markedly different across the two studies, which, in turn, renders the calculation and comparison of effect sizes across the divergent improvement criteria problematic. Nonetheless, these data certainly reveal some important trends about the efficacy of pharmacological treatments for depressed youngsters and the findings are congruent with most, if not all, of the qualitative interpretations presented within each of the studies.

Second, as a number of researchers have suggested (Anderson, 1995; Birmaher et al., 1996a; Kye & Ryan, 1995), firm conclusions about the efficacy of antidepressants in young people cannot be made until a larger body of literature is accumulated whereby some of the methodological limitations are addressed in future studies. Furthermore, of the 14 pharmacological studies included in the present analysis, only two controlled trials (14%) were located that utilized an SSRI (Emslie et al., 1997; Simeon et al., 1990). The
vast majority of the trials (n = 11) used TCAs (79%). Thus, the negative main effects of antidepressant medications for early-onset depression are attributable primarily to a lack of TCA efficacy, not necessarily pharmacotherapy overall.

Third, the number of weeks of pharmacotherapy across studies ranged from 4 to 8 weeks (median = 6). Thus, given the fact that it often takes a month or more to evidence benefits from TCAs (Duman et al., 1997), it is possible that posttreatment evaluations were premature and not enough time had passed to accurately assess the effects of the medications.

Given the limited evidence of TCA efficacy in children and the curiously high placebo-response rates in younger groups, several researchers have attempted to explain these findings based on the principles of pharmacokinetics. According to Stedman’s Medical Dictionary (1995), pharmacokinetics is defined as “the process by which a drug is absorbed, distributed, metabolized and eliminated by the body” (p. 630). For example, Geller et al. suggested that more frequent dosing is required for children and adolescents in light of the fact that they metabolize and excrete many of the antidepressant compounds more rapidly than adults (Geller et al., 1986). Similarly, several authors have reported statistically significant correlations between higher plasma levels of TCAs and clinical improvement in children with major depression (e.g., Preskorn et al., 1982; Puig-Antich et al., 1987). However, this finding is equivocal and in need of replication (Birmaher et al., 1996a). For instance, Clein and Riddle (1995) reported that the lack of demonstrated efficacy of TCAs in double-blind, placebo-controlled studies in the treatment of childhood depression does not appear
to be a pharmacokinetic phenomenon because frequent dosing has adequately addressed this issue. (p. 68)

Other pharmacokinetic hypotheses have been proposed to explain the lack of antidepressant efficacy in children and adolescents. Murrin, Gibbens, and Ferrer (1985) suggested that because most of the antidepressant medications (i.e., tertiary amines, noradrenergic TCAs) that were used in the trials purportedly enhance noradrenergic functioning, the fact that the noradrenergic systems are not fully developed until adulthood might dampen or limit the effectiveness of such medications in children. In contrast, there is evidence that the serotonergic and cholinergic systems develop earlier (Clein & Riddle, 1995), which might set the stage for additional controlled studies of the SSRIs in children and adolescents in the future. It has also been suggested that the lack of positive response to antidepressant medications in adolescents might be due to hormonal factors that interfere with certain TCAs during puberty (Birmaher et al., 1996a).

In terms of the high placebo response rates in child and adolescent samples, a number of hypotheses have been advanced. For example, Morris and Beck (1974) suggested that children have generally higher placebo response rates than adults. In addition, Birmaher et al. noted a number of potential factors associated with the high placebo response rates including: (a) the instability of affective states in juvenile groups; (b) the inclusion of children who evidence only mild to moderate depressive symptomatology; and (c) the prevalence of comorbid syndromes, especially disruptive behavioral disorders (Birmaher et al., 1996a).

A third and final comment about the interpretation of the main effects of
pharmacological trials is based on the interaction between study variables (e.g., length of treatment, age, treatment setting, medication type) and outcome. Indeed, a number of interesting trends were revealed about the efficacy of pharmacological interventions when several variables were taken into account. These data are summarized in the following section on Study Variables.

Summary

In summary, it appears that psychosocial interventions lead to substantial reductions of depression in youngsters, whereas pharmacological interventions are not superior to pill placebo in ameliorating depressive sequelae in children. There was only one study (Mandoki et al., 1997) in which pharmacotherapy was compared with an active psychosocial treatment and both interventions led to substantial reductions in depression as determined by the outcome measures. Nonetheless, because the evidence is somewhat indirect, it appears that psychosocial interventions are superior to pharmacotherapy in treating depression in youth given the current findings reported in the literature. At the same time, the methodological limitations of the pharmacological studies are noteworthy and may be masking otherwise efficacious findings. Further, in light of the interactions between study variables and outcomes that are discussed in the following section, comparisons between psychosocial and pharmacological interventions are placed within a more appropriate context.
Study Variables

Age

As described previously, a false dichotomy was created whereby child studies were defined as having a mean age of 12 or younger and adolescent studies were characterized by a sample with a mean age of 13 or older. In the psychosocial studies, the mean effect size for adolescent samples were higher than the mean effect size for child samples, regardless of experimental design. These findings were consistent with a prior study in which the mean effect size was higher for adolescents than for younger children (Weisz, Weiss, et al., 1995). However, these data were divergent from an earlier study in which psychotherapy was more efficacious for children than for adolescents (Weisz et al., 1987).

Similarly, the interaction between age and pharmacotherapy indicated that medication was more effective for adolescents than for children. Indeed, the mean ES was almost twice as large for adolescents (0.28) as it was for children (0.15). Thus, overall, it appears that psychosocial and pharmacological interventions for depressed youth are more effective for adolescents than for children. Although the reasons for these moderate findings of differential effectiveness based on age are difficult to determine, some possible hypotheses are offered. Based on the descriptive data, the modal psychosocial treatment was cognitive-behavioral therapy in group, individual, or family formats. Further, given the fact that a number of CBT interventions might depend upon a certain level of cognitive development, younger children might not be as
responsive to more cognitively sophisticated therapies. Thus, the overall findings, which consist of a predominance of CBT interventions, might favor older subjects with the assumption that they have better developed cognitive skills and derive increased benefits from the cognitively oriented interventions. With regard to the modest differential effects based on age in pharmacological studies, these findings might be attributable, in part, to the predominance of TCA trials and the empirical data that indicate younger children may not respond favorably to TCAs since their noradrenergic systems are not fully developed until late adolescence and early adulthood (Murrin et al., 1985).

Sex

Psychosocial interventions for depression were somewhat more efficacious when the average percentage of female subjects was 60% or greater. These findings are consistent with the data from other reviews that lend support to the trend of differential efficacy in favor of female subjects (Casey & Berman, 1985; Weisz, Weiss, et al., 1995). However, these findings were far from compelling and more information is needed to better understand the relationship between sex and the treatment of early-onset depression, especially given the epidemiological data, which indicate that the female-to-male ratios for depressive illness begin to approximate the base rates for adult depressive disorders (i.e., approximately 2:1) during middle to late adolescence. A possible explanation for the differential outcomes in favor of female subjects might be related to societal expectations regarding gender-appropriate behavior. That is, given the fact that females are, in general, socialized and expected to be more emotionally expressive than
their male counterparts, most forms of psychotherapy in Western culture are congruent with these socialization forces in that they promote and reinforce increased levels emotional expression. As such, female subjects potentially have an advantage over the male subjects because many of our therapeutic regimens are consistent with what they have been socialized to do from a very early age.

Referral Source

The data on the interaction between referral source and treatment outcome revealed some interesting trends. In controlled studies, 18 of the 21 (86%) interventions included nonreferred subjects and yielded an ES of 0.79. Because this ES estimate encompassed the vast majority of controlled treatments, it closely approximates the overall ES for controlled studies (0.74). Although three of the controlled studies included referred subjects and produced a smaller ES (0.44), this finding was not robust and included two outlier ES estimates near zero. Similarly, the majority of pre-/posttreatments (80%) included clinically referred subjects and produced a mean effect size of 1.20, which is commensurate with the overall ES for within-group designs (1.14). Further, although the mean effect size for nonreferred subjects was somewhat smaller than the mean effect size for referred subjects, the finding was based on only three studies so it must considered tentative. In light of these data, it appears that psychosocial interventions led to substantial reductions in depressive sequelae regardless of whether the sample was clinically referred or not. In sum, these findings suggest that analog studies can be used to provide useful data in the treatment of clinically depressed youth.
Treatment Setting

The school \((n = 15)\) was the modal setting (75%) for controlled psychosocial treatments and yielded an average ES (0.75), which was equivalent with the overall effect size for between-group designs (0.74). Similarly, the vast majority of pre-/posttreatments took place in outpatient clinics \((n = 12; 80\%)\) and produced a mean effect size (1.22) that was roughly commensurate with the overall pre-/post-effect size (1.14). Although there were only two inpatient psychosocial treatments, they were substantially less effective (0.29) than interventions conducted in schools or outpatient clinics. A similar pattern of differential efficacy was apparent in the pharmacological trials as well. Medication trials conducted in inpatient settings \((n = 6)\) were essentially ineffective (effect size = -0.06) when compared to outpatient pharmacotherapy \((n = 8; \text{effect size} = 0.26)\).

Overall, when the treatment setting and the referral source are taken into account, controlled psychosocial interventions involved mostly nonreferred subjects in school settings and produced moderately large treatment effects, whereas pre-/posttreatments included a predominance of clinically referred youngsters in outpatient settings and also led to substantial reductions in depressive sequelae. Further, it appears that inpatient treatments for depression are substantially less effective than interventions conducted in other settings, regardless of whether they are psychosocial or pharmacological in nature. This finding might be associated with the notion that inpatient subjects are not only clinically referred, but perhaps more seriously depressed than their peers who are receiving treatment in schools or outpatient clinics (Wiesz, Donenberg, et al., 1995). In addition, subjects from inpatient settings often present with a broader range of risk factors than outpatient groups (e.g., multiple diagnoses, legal and family problems). Thus,
treatments focused on a specific problem (i.e., depression) may not be adequate for children who present with more complicated sequelae.

**Type of Treatment**

Cognitive-behavioral therapy (CBT) was, by far, the most common type of psychosocial treatment for depressed youth regardless of experimental design or format (individual, group, family). Of the 36 treatments reviewed, 17 (47%) were based on cognitive-behavioral principles and 11 (31%) were cognitive-behavioral group interventions. Despite the fact that the CBT group was the modal controlled treatment, it was not the most effective despite a moderately large and robust effect size of 0.69. Lewinsohn et al. (1990) combined CBT group with family therapy which produced a mean effect size of 1.45. In addition, two group relaxation regimens (Kahn et al., 1990; Reynolds & Coats, 1986) were found to be efficacious in reducing depressive symptoms in children when compared to no-treatment control group subjects (effect size = 1.43). Other controlled interventions that yielded equivalent or more substantial effect sizes than CBT group included social skills group therapy ($n = 1$; effect size = 0.73) and behavioral problem-solving ($n = 1$; effect size = 0.68). Thus, although CBT interventions are frequently implemented and widely regarded as significantly superior to other interventions (e.g., Jayson, Wood, Kroll, Fraser, & Harrington, 1998), this assertion might be an overstatement and it is not necessarily supported by the existing empirical data. At the same time, there were a limited number of non-CBT studies, so the stability of these findings remains in question.

In terms of the 15 treatments examined in pre-/poststudies, although CBT
interventions were common (n = 4; 27%), there was balance and diversity among the remaining treatments including three social skills groups (20%), three nondirective individual interventions (20%), interpersonal therapy (n = 1; 7%), individual relaxation (n = 1; 7%), structured behavior family therapy (n = 1; 7%), and role play therapy (n = 1; 7%). Similar to the findings from controlled treatments, CBT interventions, while efficacious, were not vastly superior to a number of other types of psychosocial interventions. Indeed, interpersonal therapy, structured behavioral family therapy, and nondirective interventions all produced large and roughly commensurate treatment effects.

Another noteworthy finding regarding treatment type and effect size was the apparent detrimental impact of attention-placebo conditions. Two studies (Butler et al., 1980; Liddle & Spence, 1990) included attention-placebo groups that produced a detrimental effect (-0.83) when compared to no-treatment controls. This was a curious finding given the intended purpose of an attention-placebo condition, which purportedly controls for improvement related to variables such as equal time and attention. However, in the case of these two attention-placebo conditions, simply measuring levels of depression (i.e., no-treatment controls) was more efficacious than providing a sample of children with nonspecific time and attention.

In regards to the efficacy of particular types of medications used in the pharmacological trials, amitriptyline (Elavil) produced the highest ES (0.40; n = 3), followed by desipramine (Norpramin; 0.39; n = 2) and fluoxetine (Prozac; 0.30; n = 2). The least effective medications were imipramine (Tofranil; 0.02; n = 4) and nortriptyline
(Pamelor; -0.15; \( n = 2 \)). As mentioned previously, the predominant type of medication used in child pharmacological trials was TCAs, which have not been shown to be effective in younger groups. On the other hand, the most recent and large controlled medication trial (Emslie et al., 1997) utilized an SSRI (i.e., Prozac) that proved to be substantially superior to pill placebo in ameliorating early-onset depression (Mean effect size = 0.59). As such, this relatively efficacious finding is concealed by the overall effect size for pharmacological studies.

In sum, the interaction between treatment type and ES reveals some important findings. For instance, it appears that although CBT produces large and robust treatment effects, other interventions lead to equally substantial improvements in depression. This is not necessarily an endorsement of the dodo verdict in terms of early-onset depression treatment outcomes, but the data do not support the assertion that one type of treatment, namely cognitive-behavioral therapy, is significantly superior to other types of interventions. In regards to pharmacological medications, Elavil, Norpramin, and Prozac appear to be moderately superior to placebo in treating child and adolescent depression and there is developing evidence that SSRIs might be the medication of choice in future clinical trials.

Further, given the evidence that SSRIs are potentially less dangerous and produce fewer side effects than TCAs, these data provide some compelling and encouraging reasons to substantially increase our empirical understanding of the efficacy of such medications for the treatment of depressed youth.
Other Treatment Variables

The data regarding the interaction of frequency (i.e., number of sessions) and treatment outcome for depressed youth revealed some interesting trends. For example, as mentioned above, the length of the pharmacological trials and the TCA studies in particular was perhaps too brief (i.e., 4-8 weeks) to detect treatment effects, given that it often takes up to a month to evidence benefits. In addition, there was a modest negative relationship between the length of therapy (i.e., number of weeks) and the mean effect size \( r = -0.40; \ p = .07 \). That is, as the number of treatment weeks increased, the average effect size decreased. As such, these data might indicate that the treatment effects level off or diminish over time, but these findings should be interpreted with caution given the fact that they were not statistically significant nor practically compelling. Further, this is an important empirical question given the evidence from the adult literature, which indicates that most of the benefits of psychotherapy are obtained during the first 8-10 sessions. However, this does not necessarily imply that important gains are not made after the first 10 sessions. Indeed, in a number of health-related treatments (e.g., chemotherapy for cancer), the patient gets substantially worse before getting better. Thus, for individuals who suffer from some forms of depression, it is possible that the course of the disease worsens before improving, and discontinuing treatment prematurely based on inadequate data would be a disservice to clients. In sum, more information is needed to better understand the relationship between these variables and treatment efficacy for child and adolescent depression, especially given the current economic climate in which brief, empirically supported treatments are the battle cry of many, if not all, insurance companies and managed-care organizations.
Assessment Measures

As described previously, measuring treatment efficacy depends exclusively on assessment technology. While there were a number of studies that employed multiple depression measures to determine outcome (e.g., Brent et al., 1997; Emslie et al., 1997; Kahn et al., 1990), there was a general overreliance on self-report measures across psychosocial and pharmacological studies. Thus, given the evidence that depression, as measured by self-report devices, attenuates over time and repeated administrations, it is likely that some of the improvement depression treatment studies was attributable to extraneous variables such as the mere passage of time and regression towards the mean (Hsu, 1995). These data highlight the importance of no-treatment control groups and a broad array of assessment devices in clinical trials.

In addition, every intervention study involves demand characteristics. For example, when children present for treatment, they are aware that others expect them to improve and might therefore self-report fewer symptoms in an effort to confirm that expectation (Orne, 1969; Rosenthal, 1995). It is plausible that older relative to younger participants and female compared to male respondents are more sensitive to the demand characteristics of the treatment and testing situations. Thus, the larger effect size estimates for older subjects and female majority samples may reflect expectancy effects rather than genuine effects of treatment.

Furthermore, practical comparisons of two different measures within the same study indicated that some instruments revealed moderately divergent effect sizes. For
instance, the effect size difference between two self-report measures of depression in one study (Kahn et al., 1990) was substantial (0.65). In other examples, there were discrepant findings within the same study depending on the source of the outcome data. Using the percent improved data, the ES in the study by Geller et al. (1990) was -0.61 as compared to the SMDES on the CDRS (0.49). Moreover, in the study by Geller et al. (1992), the effect size based on the percent improved was 0.46 as compared to a less impressive effect size on the CDRS of 0.12. Thus, these data illustrate rather clearly that it can be dangerous to make firm decisions about treatment efficacy given the limitations of assessment procedures. On the other hand, effect size estimates based on two different measures were remarkably convergent in another study (e.g., 0.04 difference between BDI and RADS; Reynolds & Coats, 1986). Overall, it appears that the best way to prevent a great deal of measurement error in child depression outcome studies is to administer a broad array of devices that purport to measure depression and related constructs.

**Therapist Variables**

Professional therapists consistently produced a higher mean effect size in treating depressed youth than either graduate trainees or paraprofessional therapists, regardless of the experimental design. Graduate trainees, in turn, were more effective than their paraprofessional counterparts who produced the smallest effect sizes overall. These data are congruent with the findings from large meta-analytic reviews which indicate that professional therapists are more effective than their less trained counterparts (Stein & Lambert, 1995; Weisz et al., 1987).
Similar to the data regarding referral source and treatment setting, the majority of controlled treatments (with nonreferred subjects in schools) were most often led by graduate trainees. In contrast, pre-/posttreatments (with referred subjects in outpatient clinics) were led predominantly by professional therapists. Nonetheless, these findings lend support to the idea that increased levels of training and experience in psychotherapy lead to enhanced treatment outcomes, much to the delight of graduate training directors everywhere.

Severity of Illness and Comorbidity

Overall, it appears that psychosocial interventions for depressed youth were effective regardless of whether the children were treated based on depressive symptomatology or a diagnosis of major depressive disorder and/or dysthyrnia. In contrast, the mean effect sizes of prevention studies were substantially lower across controlled and pre-/posttreatment designs. This finding was expected, based on the intuitive assumption that the children in at-risk studies were less depressed that the subjects in symptom- or diagnosis-based trials. Further, in prevention studies, the failure to detect elevated levels of depression is not only expected, but it is an artifact of measurement limitations (i.e., floor effect, limited sensitivity to low levels of depression). The lack of treatment effect in prevention studies does not diminish their importance. Indeed, it is difficult to assess when something did not happen! At the same time, it illustrates the need to follow at-risk samples longitudinally in comparison to no-treatment controls.
Unfortunately, as mentioned previously, the data regarding comorbidity were limited. This was especially disappointing given the evidence that comorbid conditions involving depression lead to poorer outcomes, prolonged morbidity, and increased vulnerability to future bouts of depression (e.g., Kovacs, 1998; Newman et al., 1996). However, there was one provocative finding regarding the interaction of comorbidity and depression treatment outcome. In a pharmacological trial in which the efficacy of imipramine was examined, Hughes, Preskorn, Wrona, et al. (1990) reported that patients with major depressive disorder and comorbid anxiety disorders responded much better to the medication (mean effect size = 1.02) than those with major depressive disorder and comorbid conduct disorder and/or oppositional defiant disorder (mean effect size = -0.88. Thus, while there appeared to be null main effects, the data clearly revealed evidence of differential efficacy when comorbid diagnoses were taken into account. Further research is needed to better understand the relationship between comorbid diagnoses and the treatment of child and adolescent depression.

Quality of Study

There is a moderately strong and statistically significant relationship between the quality of study and the mean effect sizes for early-onset depression outcome studies. This finding was contrary to the hypothesis proposed by Wilson and Rachman (1983), wherein it was suggested that methodologically weaker studies might skew the results from meta-analytic reviews in the direction of an overestimation of treatment effects. However, the data from the present study were congruent with the findings reported by Weiss and Wiesz
Weiss and Wiesz (1990) whereby increased levels of methodological rigor were associated with higher mean effect size estimates. Further, Weiss and Wiesz suggested that experimental factors accounted for almost as much variance as the more substantive elements such as treatment type and age. Therefore, these data provide compelling evidence that methodological factors should be carefully considered during the development and implementation of experimental treatments for child and adolescent depression. At the same time, it appears less likely that studies with more limited controls will dangerously skew the findings.

Summary and Clinical Implications

The overall findings of this meta-analysis indicate that several different psychosocial interventions for early-onset depression produce moderate to large treatment gains that are clinically meaningful for many afflicted youth. Further, it appears that psychosocial treatments are, in general, superior to pharmacological regimens in treating depressed children and adolescents. However, there is also recent evidence that SSRIs such as fluoxetine are efficacious, and they will likely play an increased role in the management of affective illness in youngsters.

In terms of the clinical implications of these findings, it appears that two related but unique bodies of literature within the domain of child depression treatment studies have developed. On the one hand, there are a number of controlled studies with nonreferred children suffering from depressive symptoms who derive substantial benefit from school-based interventions led mostly by graduate students. On the other hand,
there are a large number of pre-/poststudies with referred children suffering from depressive diagnoses who benefit greatly from outpatient treatments led predominantly by professional therapists. Taken together, the implications are promising and likely generalizable to everyday clinical practice since the findings are not simply based on analogue samples or mildly depressed youngsters. Further, it appears that clinicians can select from and even combine various treatment technologies that produce roughly commensurate effects in ameliorating depressive sequelae in their clients. Finally, given the data regarding the relative superiority of psychosocial interventions over medication for depressed youth, clinicians should consider psychosocial treatments to be first-line interventions until there are more definitive answers regarding the efficacy and safety of pharmacotherapy for child and adolescent depression. This may contradict the oft-held assumption that medication is the treatment of choice when compared to psychosocial alternatives to treating psychopathological disorders. However, based on the findings from this review, this assumption should be challenged and the present data should be integrated into clinical practice, regardless of whether one is a pediatrician, psychologist, psychopharmacologist, or psychiatrist. Furthermore, while it is true that recent results from pharmacological trials have showed increased efficacy (e.g., Emslie et al., 1997), one should remain mindful of the fact that psychosocial interventions produce moderate to large treatment effects without the possibility of noxious side effects or an adverse impact on child development.
Limitations and Future Research

Perhaps the most compelling limitation of the present investigation is the fact that the vast majority of the 37 psychosocial and pharmacological studies (29; 78%) were published after 1990. Thus, this is a relatively new and quickly expanding body of literature. As such, it was a challenge to keep up with all of the new developments regarding the treatment of early-onset depression and it is likely that some important findings were not included in the present analysis (i.e., file drawer problem, unpublished manuscripts, etc.). Another limitation is based on the calculation of ES for pharmacological trials. As mentioned previously, the pharmacological findings represent a synthesis of potentially flawed data wherein effect sizes were calculated based on a transformation of divergent dichotomous variables (i.e., improvement in the active and placebo conditions).

Given the overall findings and these limitations, there are several important recommendations regarding future empirical efforts to investigate the efficacy of treatments for early-onset depression. First, future researchers should explicitly report findings based on age, sex, and comorbid conditions, in light of the limited interactional data that could be obtained regarding these important variables. Second, more consistent improvement criteria should be developed for pharmacological trials to facilitate more useful comparisons across studies. Third, researchers should attend closely to methodological variables when developing future outcome trials given the positive correlation between experimental rigor and effect size. Finally, similar to adult studies,
large controlled clinical trials examining the efficacy of psychosocial and pharmacological interventions for child and adolescent depression should be developed and implemented to provide additional insights into this vital health-care issue.

In terms of how these findings might inform policy, a few recommendations are offered. First, given the evidence that child and adolescent depression is prevalent and leads to a number of negative outcomes, diligent efforts are needed to fund and develop effective treatments for these disabling conditions. Second, more information regarding the prevention, identification, and treatment of early-onset depression should be disseminated to a broad range of health-care providers (e.g., pediatricians), child service professionals (e.g., day-care providers), and child-focused settings (e.g., schools). Finally, the treatment of early-onset depression should be considered a vital health-care issue worthy of considerable time and expense in hopes of easing the suffering of a substantial number of young people.
REFERENCES


APPENDIX

Coding Form

<table>
<thead>
<tr>
<th>Column(s)</th>
<th>ES1</th>
<th>ES2</th>
<th>ES3</th>
<th>ES4</th>
<th>ES5</th>
<th>ES6</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Study ID #</td>
</tr>
<tr>
<td>3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ES #</td>
</tr>
<tr>
<td>5,6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Year of Publication</td>
</tr>
<tr>
<td>7-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean age at beginning</td>
</tr>
<tr>
<td>11-13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% female</td>
</tr>
<tr>
<td>14-16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total n Experimental</td>
</tr>
<tr>
<td>17-19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total n Control</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Referral: 1= clinical                       2=nonclinical</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Setting: 1=school; 2=outpt; 3=inpt</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Therapist: 1=paraprofessional; 2=graduate; 3=professional</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Design: 1=pre/post; 2=exp (i.e., control group)</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Comorbid Dx: 1=yes; 2=no</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Type:</td>
</tr>
<tr>
<td>26,27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Level of intervention: 1=prevention</td>
</tr>
<tr>
<td>28-30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2=symptomatology</td>
</tr>
<tr>
<td>31-33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3=diagnosis</td>
</tr>
<tr>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total # of sessions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total # of weeks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total # of hours</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Type of intervention: 1=ind; 2=group; 3=family</td>
</tr>
<tr>
<td></td>
<td>Treatment modality:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>CBT; 2=RP; 3=Relax; 4=SS;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Psyed; 6=NDST; 7=IPT; 8=fam;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>UC; 10=Con; 11=plac;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>medication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Self-Report Measure:</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>CDI; 2=BDI; 3=RCDS;</td>
</tr>
<tr>
<td>4</td>
<td>HRSD; 5=CESD; 6=RADS;</td>
</tr>
<tr>
<td>7</td>
<td>MFQC; other (specify)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Dx Interview: 1=yes; 2=no</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Other measures: 1=yes; 2=no</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SMDES Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>write in...</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SMDES ±_ _ _</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Follow-up: 1=yes; 2=no</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Length of Follow-up (wks)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Follow-up SMDES</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Quality of Study: Validity:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>unacceptable</td>
</tr>
<tr>
<td>2</td>
<td>poor</td>
</tr>
<tr>
<td>3</td>
<td>fair</td>
</tr>
<tr>
<td>4</td>
<td>good</td>
</tr>
<tr>
<td>5</td>
<td>excellent</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Threats to Validity (0-3):</th>
</tr>
</thead>
<tbody>
<tr>
<td>maturation:</td>
<td></td>
</tr>
<tr>
<td>history:</td>
<td></td>
</tr>
<tr>
<td>testing:</td>
<td></td>
</tr>
<tr>
<td>instrumentation:</td>
<td></td>
</tr>
<tr>
<td>regression:</td>
<td></td>
</tr>
<tr>
<td>selection bias:</td>
<td></td>
</tr>
<tr>
<td>experimental mortality:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0=not a plausible threat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>minor threat</td>
</tr>
<tr>
<td>2</td>
<td>plausible threat</td>
</tr>
<tr>
<td>3</td>
<td>by itself could explain findings</td>
</tr>
</tbody>
</table>

Note: Aer Exer=Aerobic Exercise; Alt-Tx=Alternative Treatment; Attn-Plac=Attention-Placebo Group; CBT-G=Cognitive-Behavioral Therapy-Group; CBT-I=Cognitive-Behavioral Therapy-Individual; Consult=Consultation; F=Family Therapy; IPT-I=Interpersonal Psychotherapy-Individual; NFI-I=Non-focused Intervention-Individual; NST-I=Non-Directive Supportive Therapy-Individual; P=Parent; PS-G=Problem-Solving Group; PsyEd-G=Psychoeducational Group; Rlx-I=Individual Relaxation Therapy; Rlx-G=Group Relaxation Therapy; RP-G=Role-Play Group; SS-G=Social Skills Group; SBFT=Structured Behavior Family Therapy.
VITA

Kurt David Michael

Duke University Medical Center
Division of Medical Psychology
P.O. Box 3362
Durham, NC 27710

4304 Waterford Valley Drive, #1821
Durham, North Carolina 27713
H: (919) 544-7980
W: (919) 416-7211
e-mail: kmichael@duke.edu

Education and Training


Academic Awards

1996 - 1997  Dale and Adele Young Scholar, Utah State University.
1997  Phi Kappa Phi, National Honor Society.
1995 - 1996  Rural Psychology Training Grant, Utah State University.
1988  Departmental Honors in Psychology, University of Colorado at Boulder.
1988  Excellence in Research, Rocky Mountain Psychological Association.
1987  Mortar Board, National Honor Society for Academics and Community Service.
1986  Psi Chi, National Honor Society in Psychology.
Service and Affiliations

1996 - 1997  Graduate Student Representative, Combined Psychology Program, Utah State University.
1995, 1996  Graduate Student Representative, Clinical Psychology Faculty Search Committee.
1992  Member, National Advisory Board for Victims of Drug Related Crimes.
1986 - present  Graduate Affiliate, American Psychological Association.

Clinical Experience

Current Clinical Position

Clinical responsibilities on the following rotations:
  Pediatric Psychology- assessment of psychological functioning associated with medical sequelae (e.g., brain tumors, seizure disorders, learning disabilities, sickle-cell disease, developmental disorders), child therapy, & consultation.
  Durham Community Guidance Clinic- individual and group psychotherapy, assessment, behavioral consultation, & community-based intervention.
  Adolescent Inpatient Unit- diagnostic interviewing, assessment, & group therapy.
  North Carolina Center for Child and Family Health- assessment and treatment of children and families who have experienced traumatic events.
  Family Studies Clinic- family and marital psychotherapy, & parent training.

Previous Clinical Positions

1997 - 1998  Graduate Assistant, University Counseling Center, Utah State University, Logan, Utah.
Provided individual, couples, and group therapy under the supervision of licensed clinical psychologists. Conducted intake assessments.
Hours: 1000.

1997 - 1998  Group Co-Therapist, Psychology Community Clinic, Utah State University.
Provided weekly group psychotherapy for survivors of child sexual abuse.
Direct Hours: 100.

1996 - 1997  Psychology Specialist, Clinical Services, Center for Persons With Disabilities, Utah State University.
Conducted comprehensive child evaluations. Provided behavioral consultation and psychological treatment to children and families.
Supervised practicum students and facilitated social skills groups.
Hours: 1250.

1995 - 1997  Therapist, Psychology Community Clinic, Utah State University.
Provided individual, couples, and family psychotherapy under the supervision of licensed clinical psychologists. Responsibilities included: diagnostic interviews, report writing, and case management. Direct Hours: 230.

Conducted child psychoeducational, behavioral, and psychological assessments. Direct Hours: 320.
1994 - 1995

**Psychoeducational/Mental Health Specialist**, Community Family Partnership, Center for Persons with Disabilities, Utah State University. Conducted preschool, developmental, and psychological assessments and provided child and adult mental health therapy. Hours: 1000.

1993 - 1994

**Clinical Supervisor**, Substance Abuse Unit, Beaumont Juvenile Correctional Center, Virginia Department of Youth and Family Services, Beaumont, Virginia.
Provided clinical supervision to counselors working with incarcerated juvenile offenders with a history of antisocial behaviors, aggression, substance abuse, and drug dealing. Facilitated therapy groups, supervision groups, and treatment team meetings. Responsible for program development, implementation, & management. Hours: 2610.

1991 - 1993

**Counselor**, United Methodist Family Services, Richmond, Virginia.
Provided counseling and supervision to emotionally disturbed, adjudicated, and chemically dependent children in a long-term residential treatment milieu. Led extended hiking, rock climbing, and camping trips. Facilitated groups and high and low ropes course events. Hours: 4300.

1992 - 1993

**Chemical Dependency Counselor/Psychiatric Technician**, Charter Westbrook Hospital, Richmond, Virginia.
Facilitated chemical dependency treatment groups on the child and adolescent acute and long-term inpatient units. Supervised patients and served as a member of an interdisciplinary treatment team. Direct Hours: 400.

1990 - 1991

**Senior Counselor**, Therapeutic Community, Shalom et Benedictus, Stephenson, Virginia.
Provided counseling to chemically dependent, adjudicated, and emotionally disturbed adolescents. Facilitated therapy groups and treatment team meetings. Responsible for clinical supervision and staff scheduling. Assisted the Program Director in the daily management of the therapeutic milieu. Hours: 3800.

1989 - 1990

**Community Assistance Coordinator**, Shalom et Benedictus, Stephenson, Virginia.
Provided counseling and consultation to children and their families. Facilitated support groups in city and county schools such as: children of divorce, substance abuse aftercare, and social skills. Hours: 2000.

1989 - 1991

**Counseling Consultant**, Shenandoah University Counseling Clinic, Winchester, Virginia.
Provided individual and group counseling for students. Served as a crisis interventionist and administered substance abuse assessments. Direct Hours: 600.

1986 - 1989

**Student Director**, Rapline, Wardenburg Student Health Center, University of Colorado at Boulder.
Served as a crisis interventionist trained to deal with suicide, domestic violence, rape, depression, and other conditions. Responsibilities included: recruitment, training, and supervision of phone counselors. Direct Hours: 400.
Clinical (Year-Long) Practica

1997 - 1998  Behavioral Health Unit, Logan Regional Hospital, Logan, Utah. Conducted psychological evaluations, wrote psychological reports, and co-facilitated group therapy sessions on an inpatient psychiatric unit. Participated in interdisciplinary treatment team meetings.

1997 (Summer)  Psychology Community Clinic, Department of Psychology, Utah State University, Logan, Utah. Provided individual, couples, and family psychotherapy under the supervision of licensed clinical psychologists. Wrote psychological reports.

1996 - 1997  Utah State University Counseling Center, Logan, Utah. Provided individual, couples, and group therapy under the supervision of licensed clinical psychologists. Conducted intake assessments.

1995 - 1996  Ogden City School District, Ogden, Utah. Conducted comprehensive psychoeducational assessments, conducted behavioral observations, and facilitated groups for preschoolers with developmental disabilities and behavioral disorders.

1994 - 1995  Psychology Community Clinic, Department of Psychology, Utah State University, Logan, Utah. Provided individual, couples, and family psychotherapy under the supervision of licensed clinical psychologists. Wrote psychological reports.

Clinical Consultation

1995 - 1998  Psychometrician, for David Stein, PhD, Private Practitioner, Logan, Utah. Conducted intellectual, memory, and psychological assessments as part of comprehensive vocational rehabilitation/disability evaluations.

Teaching Experience

1997  Instructor, Psychology 510/610, History and Systems, Utah State University, Logan, Utah. Responsibilities included: lecture preparation and delivery, exam construction and administration, and grading.

1996  Graduate Teaching Assistant, Psychology 631, Intellectual Assessment, Utah State University, Logan, Utah. Responsibilities included: grading assessment protocols and reports, and didactic instruction.


1995 - 1996  Graduate Teaching Assistant, Psychology 101, General Psychology, Utah State University. Responsibilities included: lecture preparation and delivery, supervision of 6 lab instructors, exam construction and administration, tutoring, exam review sessions, and grading.
1993 - 1994  Adjunct Training Faculty, Virginia Department of Youth and Family Services, Richmond, Virginia. Trained Department personnel on issues such as: chemical dependency, drug trafficking, milieu treatment for juvenile offenders, adventure-based treatment, and ropes course facilitation.


1987 - 1988  Teaching Assistant, General Psychology, University of Colorado at Boulder. Responsibilities included: lecturing, tutoring, exam review sessions, and grading.

Research Experience

1997 - present  Child clinical research with Dr. Susan Crowley, Utah State University. Integrative review of the efficacy of treatments for childhood depression.

1996 - 1997  Child health psychology research with Drs. Kevin Masters and Gretchen Gimpel, Utah State University. Worked as part of research team examining a number of variables related to child health psychology such as: asthma, exercise, and nutrition.

1996  Child clinical research with Dr. Ken Merrell, Utah State University. Convergent validity study of several self-report instruments that purport to measure depression, anxiety, and other internalizing symptoms.


1988  Researcher I, Pediatric Psychiatry, National Jewish Center for Immunology and Respiratory Medicine, Denver, Colorado. Responsibilities included the collection of data via: pediatric phlebotomy, viral cultures, videotaped observation, and interviews.


1987  Research Assistant, Raimy Clinic, Clinical Psychology Department, University of Colorado at Boulder. Conducted literature reviews, entered data, and helped to develop clinical research instruments for PhD students.

1986  Research Assistant, Institute for Behavioral Genetics, Boulder, Colorado. Served as a member of a research team examining the environmental and genetic influences on varied behaviors. Responsible for data collection and data entry.
Publications


Editorial Experience

Guest Editor, Special Issue on Substance Abuse Research, Counselor, September/October, 1996.

Presentations


Other Experience

1987 Management Assistant, to the Assistant Attorney General for Personnel and Administration, United States Department of Justice, Washington, DC.Reviewed and edited the Department’s manual on employee urinanalysis procedures.

Professional Licensure and Certifications

C.S.P., Certified School Psychologist, Utah State Office of Education, Salt Lake City, Utah.
L.S.A.C., Licensed Substance Abuse Counselor (1997 - present), Utah Division of Occupational and Professional Licensing.

C.A.C., Certified Addictions Counselor (1989 - present), Utah Association of Alcoholism and Drug Abuse Counselors; (formerly) Substance Abuse Certification Alliance of Virginia.

C.C.S., Certified Clinical Supervisor (1993 - present), Utah Association of Alcoholism and Drug Abuse Counselors; (formerly) SACAVA.
C.S.A.C., Certified Substance Abuse Counselor (1994 - present), Board of Professional Counselors and Marriage and Family Therapists, Virginia Department of Health Professions.

Selected Conferences and Workshops Attended

*An Afternoon with a Master*, Irvin Yalom, MD, University Counseling Center, Utah State University, April, 1998.

*Anxiety Disorders in Youth: Interventions for Children and Families*, Thomas Ollendick, PhD, Seventh Annual Virginia Beach Conference Workshop, September, 1997.

*Raising Children in a Socially Toxic Environment*, James Garbarino, PhD, Utah State University, January, 1997.

*Inside the Criminal Mind*, Stanton Samenow, PhD, Benchmark Health Systems, Salt Lake City, Utah, May, 1996.

*Exploring the Power of Group*, Robert Weber, PhD, University Counseling Center, Utah State University, April, 1996.

References

Susan Crowley, PhD  
Associate Professor, Director of Clinical Training  
Psychology Department  
2810 Old Main Hill  
Utah State University  
Logan, Utah 84322-2810  
(435) 797-1251

Melanie Bonner, PhD  
Assistant Clinical Professor  
Department of Psychiatry, Division of Medical Psychology  
Duke University Medical Center, Box 3362  
Durham, North Carolina 27710  
(919) 684-4129

Dennis Odell, MD  
Medical Director and Pediatrician  
Center for Persons with Disabilities Adjunct Associate Professor, Biology Department  
6800 Old Main Hill  
Utah State University  
Logan, Utah 84322-6800  
(435) 797-2750

Kenneth Merrell, PhD  
Associate Professor  
Division of Psychological and Quantitative Foundations  
N334 Lindquist Center  
The University of Iowa  
Iowa City, Iowa 52242-1569  
(319) 335-5495

David Stein, PhD  
Professor and Head  
Psychology Department  
2810 Old Main Hill  
Utah State University  
Logan, Utah 84322-2810  
(435) 797-3274

Patricia Truhn, PhD  
Adjunct Assistant Professor  
Department of Psychology  
Psychologist, Center for Persons with Disabilities  
6800 Old Main Hill  
Utah State University  
Logan, UT 84322-6800  
(435) 797-3822 or 797-2008