SCHOOL-BASED INTERVENTIONS FOR ANXIOUS AND DEPRESSED YOUTH: A META-ANALYSIS OF OUTCOMES

A dissertation submitted to the Temple University Graduate Board

in Partial Fulfillment of the Requirements for the Doctorate of Philosophy

---

By

Matthew P. Mychailyszyn

Temple University

August, 2011

Major Advisor:
Philip Kendall, Ph.D. – Temple University, Department of Psychology

Committee Members:
Rob Fauber, Ph.D. – Temple University, Department of Psychology
Ronald Brown. Ph.D. – Wayne State University, Department of Academic Affairs
Tania Giovannetti, Ph.D. – Temple University, Department of Psychology
Lauren Alloy, Ph.D. – Temple University, Department of Psychology
Brian Daly, Ph.D. – Drexel University, Department of Psychology

ABSTRACT

Objective: Conduct a meta-analysis of school-based interventions for anxious and depressed youth to evaluate the findings according to QUORUM guidelines. **Method:** A search of the literature was conducted using PubMed, PsycINFO, and manual searches, supplemented by contact with leading researchers. The present meta-analysis used 63 studies and investigated 8,225 participants receiving CBT and 6,986 participants in comparison conditions. **Results:** Mean pre-post effect size estimates indicate that anxiety-focused school-based CBT was moderately effective at improving symptomatology of anxiety (Hedge's g = 0.501) and depression-focused school-based CBT was mildly effective at improving depression symptomatology (Hedge's g = 0.298) for youth receiving active interventions as compared to those in anxiety intervention control conditions (Hedge's g = 0.193) and depression intervention controls (Hedge's g = 0.091). Moderators and mediators of treatment outcome were explored. **Conclusions:** School-based CBT interventions for youth anxiety and for youth depression hold considerable promise, though further investigation is still needed to identify features that optimize service delivery and outcome.

ACKNOWLEDGEMENTS

The completion of this project would not have been possible without the help and support of a number of people to whom I am eternally grateful.

First, I would like to thank my family. To my mother, for instilling in me a passion for learning. To my father, for the example he set through the sacrifices he made for our family that allowed my education to be possible. To my sister, who always supported me and whose presence was an immense comfort when I began graduate school in Philadelphia. Thank you all for shaping me into the person I am today.

Next, I would like to thank the faculty at Temple University, especially those of whom are members of my committee. Each of you has played a number of important roles over the course of my training and professional development – as teachers, supervisors, advocates, mentors, colleagues, supporters, advisors, and friends. My career in the field of psychology will forever be grounded in what I have learned from you, and for this I am truly grateful. I would also like to thank Douglas Brodman and Kendra Read for their coding assistance on this project.

Finally, and most importantly, I would like to thank my wife, Katie Mychailyszyn, who has been an angel in my life since we first met at Loyola College in Maryland. The decision to move to Philadelphia for graduate school was one of the most difficult of my life. With my education requiring that four years of our relationship be spent in different states – including the first year of our married life together – it was only with her immense patience and support that I have now reached this place. Words cannot express the depth of my gratitude for the love, friendship, and commitment she has given me over the years.

To all those whose presence has been a blessing in my life – Thank You.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Children and adolescents face many challenges: Increasing educational demands (Harris, 2008), rising rates of divorce (Heckel, Clarke, Barry, McCarthy, & Selikowitz, 2009; Messner, Bjarnason, Raffalovich, & Robinson, 2006), and media exposure to messages of violence and terrorism (Comer, Furr, Beidas, Weiner, & Kendall, 2008; Comer & Kendall, 2007), among others. These challenges contribute to youth's vulnerability to a wide range of associated mental health difficulties.

Youth psychopathology prevalence rates have been found to range from 1%-51%, with the most reliable estimates suggesting that between 12%-20% of youth are struggling with clinical-level symptoms of disorder at any given time (Roberts, Attkisson & Rosenblatt, 1998; Costello, Egger, & Angold, 2005; Tolan & Dodge, 2005). Such figures are consistent with estimates made by the United States Congress, suggesting that between 5.6 million and 6.8 million (18%-22%) youth are in need of mental health services (U.S. Public Health Service, 2000). With specific reference to anxiety disorders, prevalence rates in children and adolescents have been found to range from 2%-27% (Costello, Egger, & Angold, 2004). Similar rates – between 7% and 28% – have been found for depression, (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Lewinsohn & Clarke, 1999), with empirical evidence suggesting that its prevalence is increasing among children and adolescents (Hammen & Rudoloph, 2003). Indeed, research suggests that these two disorders are among the most prevalent categories of psychological problems among youth (e.g., Albano, Chorpita, & Barlow, 2003; Bernstein & Borchardt, 1991; Costello et al., 2003).

Anxiety and depression amongst youth are associated with negative outcomes and sequelae. Although anxious/withdrawn children may be perceived as less troublesome than those exhibiting hyperactive or oppositional behavior, they are nevertheless distressed and impaired (Ialongo, Edelsohn, Werthamer-Larsson, Crockett, & Kellam, 1994; Ialongo, Werthamer-Larsson, Crockett, & Pellam, 1995; Strauss, Frame, & Forehand, 1987; Wood, 2006). Anxiety disorders increase vulnerability to the development of comorbid conditions, and, if left untreated, may persist into adulthood and lead to the development of substance abuse problems (Albano et al., 2003; Kendall, Safford, Flannery-Schroeder, & Webb, 2004; Woodward & Fergusson, 2001). Depression is similarly associated with a range of impairments for youth struggling with this condition (Collins & Dozois, 2008). Though the symptoms alone – depressed mood, diminished interest in pleasurable activities, sleep disturbance, feelings of guilt and/or worthlessness – are sufficiently troublesome unto themselves, they often lead to additional problems of functioning, with evidence suggesting that even youth who have elevated but subclinical levels of depressive symptoms experience a wide range of impairment (Georgiades, Lewinsohn, Monroe, & Seeley, 2006; Gotlib, Lewinsohn, & Seeley, 1995). Further, and of major concern, depression amongst adolescents tends to be recurrent, with episodes of earlier onset having a longer duration and being a significant predictor for the experience of later episodes and adult depression (Wicks-Nelson & Israel, 2009).

Regrettably, a considerable portion of youth never receives any kind of help. Though as many as 40% of those with a psychiatric diagnosis and associated impairment may be accessing services across different sectors, only about one in five are receiving care from a specialty mental health provider (Burns et al., 1995). Less conservative estimates suggest that between two-thirds and three-quarters of those youth in need of help do not have access to appropriate care and are

thus left untreated (United States Congress, Office of Technology Assessment, 1991). This situation is especially troublesome with regard to anxious and depressive internalizing disorders; because such problems are often less visible to parents and teachers as compared to externalizing conditions (e.g. ADHD), the majority of youth struggling with anxiety and depression have never received treatment (Chavira, Stein, Bailey, & Stein, 2004; Logan & King, 2002). Perhaps most disturbing of all is that relatively little ground has been gained, as the estimated percentage of those with unmet needs has remained virtually unchanged over the past two decades (United States Congress, Office of Technology Assessment, 1986).

Despite strong support in favor of evidence-based practice (EBP) from the American Psychological Association (APA) and the American Academy of Child and Adolescent Psychiatry (AACAP), fewer still of the individuals receiving services are actually administered the type of empirically-supported treatments (ESTs) that research evidence has deemed "efficacious" (U.S. Public Health Service, 2000). Although the emphasized importance of reliance on an evidence base is beginning to lead to a greater focus on ESTs in the training of the newest generations of psychologists (Cukrowicz et al., 2005), a considerable gap remains. Indeed, McCabe (2004) claims that "practical guidelines for professional psychologists who may be interested in incorporating EBPs into their own work setting are not available" (p.571), while Schoenwald and colleagues (2008), comprising The Research Network on Youth Mental Health, state that "little is known about the nation's infrastructure for children's mental health services (CMHS), the capacity of that infrastructure to support the implementation of (ESTs), and factors affecting that capacity" (p.85).

There is a need for change with regard to the manner in which mental health services are delivered to youth, with a focus on increasing access to effective treatment for those in need. One

commonly suggested solution is to more fully incorporate mental health services into school systems. Indeed, the need for schools to play a larger role in the establishment and maintenance of emotional and psychological well-being for youth is widely noted (e.g., Weist, Evans, & Lever, 2003). The former Surgeon General's Report on Children's Mental Health in 2000 promoted "cost-effective, proactive systems of behavior support at the school level" and a strengthening of schools' capacity to be "a key link to a comprehensive, seamless system of school- and community-based identification, assessment and treatment services" (U.S. Public Health Service, 2000). The President's New Freedom Commission on Mental Health (2003) supported these sentiments, emphasizing the dynamic interplay between emotional well-being and academic success, and encouraging schools to act as partners in the mental health care of children. These sentiments have been echoed by governments around the world. For instance, in establishing its national action plan for mental health from 2006-2011, the Council of Australia Governments (COAG; http://www.coag.gov.au/reports/docs/AHMC_COAG_mental_health.pdf) pledged political and financial support to reforming mental health services and building partnerships that would allow a more effective institution of school-based prevention and early intervention programs for children and adolescents in need of care.

The problem of insufficient access to care has a long history, with a litany of exhortations for change being made by researchers and governmental policymakers alike. To what extent, then, have interventions for anxious and depressed youth actually been applied in the schools to meet this need, and what is the status of the collective set of outcomes? The answers to questions such as these are complex and multidimensional, involving research that spans numerous domains of investigational inquiry. As such, there is a need to integrate and evaluate the

available data that relate to the implementation of school-based interventions for child and adolescent anxiety and depression.

The purpose of the present review is to examine the empirical literature pertinent to school-based interventions that separately and jointly target youth anxiety and depression. The review is organized into five sections: (a) an overview of the research investigating the efficacy of CBT treatment for anxiety and depression; (b) a discussion of transportability and the dissemination of efficacious treatments to community settings; (c) a description of the school context, with consideration of the inherent advantages and drawbacks to this setting; (d) an overview of the classification system used for identifying types of intervention, including discussion of the differences in their intended purpose; and (e) a presentation of the current study, including the methods, results, discussion of the findings, and recommendations for future implementation of school-based interventions for youth anxiety and depression.

Treatment of Anxiety and Depression in Youth

With a commitment to extend mental health services to youth, the next question that must be asked is: "What services have been found to be helpful for children and adolescents?" One of the primary modes of intervention for youth anxiety and depression is cognitive-behavioral therapy (CBT; James, Soler, & Weatherall, 2005). CBT is a collaborative, problem-focused approach that seeks to address the underlying and maintaining factors of the child's distress (Kendall, 2006). The focus of treatment for both anxiety and depression is to help the child learn about the nature of the disorder so that he or she can become better at identifying and coping with the associated distress. This often involves a combination of psychoeducation about physiological responses to emotions, relaxation training for the easing of unpleasant somatic arousal, cognitive restructuring to address patterns of distorted and maladaptive thinking,

effective problem solving, and behavioral exercises to break negative cycles of avoidance and inactivity. In this way, youth are provided with the positive strategies and tools that allow them to more successfully manage symptoms of anxiety and depression.

Research has found that treatment for anxiety disorders in youth can be efficacious. A randomized clinical trial (RCT; Kendall, 1994) examined the efficacy of CBT for treating anxious youth (the *Coping Cat Program*; Kendall & Hedtke, 2006a; Kendall & Hedtke, 2006b). At posttreatment, 64% of those who received treatment – compared to only 5% of waitlist controls – no longer met criteria for their primary diagnosis, with improvements maintained at 1-year follow-up. Similar results were reported in a second RCT (Kendall et al. 1997), and treatment gains have been found to be maintained up to 7 years later (Kendall et al., 2004)

Barrett, Dadds, and Rapee (1996), in using an Australian adaptation of the Coping Cat, found support for the treatment with added advantages derived from the addition of a family anxiety management component. Results showed that 57% of those in CBT, and 84% of those in CBT+FAM no longer met diagnostic criteria, compared to only 26% of those on the waitlist. These gains had improved 1 year later. Further support for family cognitive-behavioral therapy (FCBT) was demonstrated by Kendall and colleagues (2008). Both FCBT and an individual CBT (ICBT) group were found to be superior to a family-based education, support, and attention control group (FESA). Although not superior to ICBT strictly in terms of responder status, FCBT did perform better when both of an anxious youngster's parents had an anxiety disorder.

Using a group approach, Silverman et al. (1999) investigated the efficacy of Group Cognitive-Behavior Therapy (GCBT). Compared to only 13% in the control group, 64% of those receiving GCBT were recovered at posttreatment, and that percentage increased to greater than 75% at 3-, 6-, and 12-month follow-ups. Flannery-Schroeder and Kendall (2000) compared

ICBT, GCBT, and waitlist (WL), and found that 50% of youth in the GCBT group and 73% of youth in the ICBT group no longer met criteria for their primary anxiety diagnosis (compared to 8% of the WL group). These improvements were maintained at 1-year follow-up (Flannery-Schroeder, Choudhury, & Kendall, 2005).

A recent and comprehensive evaluation of treatments for youth anxiety (Walkup et al, 2008), was a comparative study including CBT, medication (sertraline), their combination, and a pill placebo. Results indicated that 60% of youth receiving CBT were rated as very much or much improved following 12 weeks of treatment. The best outcomes were produced by the combination of CBT plus medication condition. CBT and medication each yielded comparable rates of improvement when administered by themselves, though less somatic side effects (e.g. insomnia, fatigue, restlessness) were associated with CBT than for medication.

Empirical support has also been found for the efficacy of cognitive-behavioral treatment of depression in youth. Research on this topic began earlier than for anxiety, with studies beginning mostly in the early 1980s. However, due to the comparatively lower prevalence rate of depression in prepubertal children (Kovacs, 1996) treatment studies initially focused more on older youth until later years. Building on initial evidence of efficacy for cognitive and behavioral modalities of treatment for depression with adults (McLean & Hakstian, 1979, Shaw, 1977), the early emerging studies (Lewinsohn, Clarke, Hops, & Andrews, 1990) for adolescent populations took a downward extension approach, focusing on modifying treatment for developmentally appropriate application with younger participants.

Additional evidence for the efficacy of cognitive-behavioral interventions for depression in youth has been obtained from studies targeting children of depressed parents. Aimed at treating subclinical levels of depression in order to prevent later onset of criterion-level

depressive episodes, Clarke and colleagues (2001) used a group-based format in which youth were taught strategies to counteract unrealistic negative thoughts, especially those related to having a depressed parent. Statistically and clinically significant preventative results were found for the sample of at-risk youth; however, these results are tempered by less clearly positive findings from an earlier similar investigation of prevention for offspring of depressed parents that delivered treatment through a family based format (Beardslee et al., 1993).

Finally, a recent evaluation of the long-term efficacy of various treatment modalities for youth depression (TADS Team, 2007) found increasing rates of response over time for adolescents receiving CBT following the cessation of treatment. Despite having significantly lower rates of response in the shorter term of 12 weeks, by 36 weeks CBT produced response rates equivalent to those of medication and medication combined with therapy. Moreover, adding CBT appeared to enhance the safety of the medication, as youth receiving combination treatment experienced significantly fewer suicidal events than those receiving medication alone.

Transportability of Interventions

Based on the accumulation of outcome studies similar to those described above, as well as conclusions reached by reviews (e.g., Collins & Dozois, 2008; Horowitz & Garber, 2006; James, Soler, & Weatherall, 2005; Ollendick & King, 2006), CBT meets established standards to be considered an efficacious treatment for disorders of anxiety and depression in youth and thus should be recommended as the first-line treatment of choice for youth struggling with these problems (APA Task Force on the Promotion and Dissemination of Psychological Procedures, 1995; Chambless & Hollon, 1998; Compton et al., 2004). Additionally, in reviewing the research, Evans and Weist (2004) conclude that providing interventions that do not have empirical support is likely to provide little to no benefit to students and schools. However,

questions remain regarding the potential for empirically-supported interventions to be successful when implemented in settings such as schools, where they are delivered by a variety of individuals to diverse populations (Owens & Murphy, 2004).

The source of and answers to these questions revolve primarily around issues of transportability — the degree to which evidence-based treatments work when implemented in community contexts (Schoenwald & Hoagwood, 2001). Issues revolving around this topic are not new, however: Over 15 years ago the *Journal of Consulting and Clinical Psychology* devoted a special issue to an examination of "how findings from carefully controlled studies of efficacious psychosocial interventions for children can be transported into naturalistic studies of the effectiveness of services" (Hoagwood, Hibbs, Brent, & Jensen, 1995). The notion of transportability has continually been discussed at multiple levels, surfacing in the Surgeon General's Conference on Children's Mental Health which promoted increased reliance on the use of "scientifically proven" mental health services (U.S. Public Health Service, 2000). Indeed, the issue is still before us today, as Ginsburg and colleagues (2008) point out that the challenges which continue to confront psychology are with regards to successful dissemination of empirically supported intervention strategies to community treatment settings – especially those serving youth from diverse racial and ethnic backgrounds. The difficulty, however, is found in the existence of a gap between research and service clinics such that results derived from the lab may be difficult to replicate in the community (Weisz et al., 1995).

The achievement of transportability thus requires a "bridging of the gap," which has also been referred to as "smoothing the trail" (Kendall & Beidas, 2007) and "translating science into practice" (Chorpita, 2003). What this entails is essentially a move from "efficacy" to "effectiveness" (Mufson, Dorta, Olfson, Weissman, & Hoagwood, 2004; Schoenwald &

Hoagwood, 2001) or from "research therapy" to "clinic therapy" (Weisz et al., 1995). The former (in each case) is often characterized by homogeneously comprised samples and clinicians with in-depth training in the use of manual-based treatments, with the goal of testing the intervention. On the other hand, the latter seeks to evaluate applications of such efficacious treatments in community settings that commonly lack such resources.

The consequence is often lower adherence to treatment fidelity and a drop-off in treatment effects compared to RCTs (Henggeler, Melton, Brondino, Scherer, & Hanley, 1997). Hence, a critical goal is to develop strategies to reduce these problems, with articles included in the special section of the *Journal of Consulting and Clinical Psychology* discussing how such a goal might be realized (Clarke, 1995; Weisz et al., 1995), the implications of transportability (Henggeler, Schoenwald, & Pickrel, 1995), and the use of manual-based treatment (Kendall & Southam-Gerow, 1995).

Although the prospect of transportability may incur greater expenses, the investment is worthwhile when compared to the prospect of providing ineffective services (Henggeler et al., 1997). Success on this front yields what Schoenwald and Hoagwood (2001) refer to as "street-ready" interventions – ones that can be applied in representative settings and systems. Thus, following empirical validation of efficacy, treatments can then be adapted for application in broader community settings.

The School Context

The dissemination and widespread implementation of empirically-supported treatments within school systems would mark a considerable change from the manner in which mental health services are traditionally provided to youth (Evans & Weist, 2004). Challenges exist with regard to schools' acceptance of a greater role in children's mental health (Owens & Murphy,

2004; Pincus & Friedman, 2004; Weisz et al., 1995) and many difficult questions remain, as yet, insufficiently answered. Schoenwald and Hoagwood (2001) inquire: What is the intervention? Who can implement it, under what circumstances, and to what effect? Additionally, Owens and Murphy (2004) ask: How effective are these treatments when delivered to diverse populations by mental health professionals in community settings who struggle with the added burdens of higher caseloads and fewer resources?

These are important issues that must be addressed, and such questions reflect a procedural challenge for school-based interventions. The link between children's mental health and academic success would seemingly provide a natural avenue for collaborative efforts between professionals in psychology and education (Mufson, Dorta, Olfson, Weissman, & Hoagwood, 2004), as research has demonstrated the deleterious effects of psychopathology on children and adolescents' school functioning (Fauber, Forehand, Long, Burke, & Faust, 1987; Ialongo et al., 1994, 1995; Jaycox, Stein, Paddock, Miles, Chandra, Meredith, et al., 2009*)*. However, developing and sustaining such relationships can be difficult. Schools may be wary of researchers; though studies often bring an influx of resources, they can likewise be terminated just as quickly. Additionally, research may be stymied when parents and/or school administrations have negative perceptions of research.

A particularly problematic obstacle for school-based mental health interventions can be getting teachers "on board" (Pincus & Friedman, 2004). Although it can be possible for other school personnel to carry the burden, teachers are often asked to play an active part in the delivery of school-based interventions, with related tasks including identification of eligible participants, completion of research measures, and perhaps even implementation of the program itself. However, these tasks may require considerable training, reflecting a time commitment that

competes with an already demanding schedule (Owens & Murphy, 2004). Combined with the possibility that children may need to spend time out of the classroom to participate in the intervention, it is understandable that teachers may not be enthusiastic about involvement in school-based research.

Where teachers have been "on board" and transportability efforts have been attempted the answer to the question of "Who can implement the intervention?" and "To what effect?" has led to largely mixed results. For instance, in a qualitative review of nine prevention efforts for youth depression (Andrews et al., 2002), the authors suggested that the differences in outcome between two particular intervention evaluations were at least partly due to the difference in program facilitator, whereby the one that used teachers as implementers yielded less positive results. In their meta-analysis of CBT for anxious youth, Ishikawa and colleagues (2007) claim that, "in terms of delivery settings the results suggest that CBT provided at university clinics was more effective. This may be due to researchers at universities being better able to ensure treatment and protocol fidelity" (p.170). Additionally, in their systematic review of school-based interventions for depression, Calear and Christensen (2010) found that studies using teachers as program leaders were associated with fewer significant effects and conclude that "this finding suggests that programs are more effective in the hands of mental health professionals or the program developers" (p.435). On the other hand, the results of a meta-analysis focusing on the Penn Resiliency Program (Brunwasser, Gillham, & Kim, 2009) for youth depression found that mean effect sizes for interventions implemented by community professionals (which included teachers) were not significantly different from those produced by interventions led by research teams. Further, based on a research review of universal school-based preventive interventions for depression, Spence and Shortt (2007) conclude that outcomes "do not appear to be related to the

level of mental health training of the person delivering the program" (p. 537). Given such inconsistencies, further research on the topic of intervention implementer appears to be warranted.

Given the existence of numerous obstacles, a valuable characteristic to possess with regards to the challenges in the school system is flexibility (Kendall & Beidas, 2007; Pincus & Friedman, 2004). The successful researcher is one who accepts the restrictions inherent to the setting and finds alternative solutions to the problems they present. This flexibility may involve occupying a less than ideal location and adapting the treatment length or course, while still adhering to treatment fidelity (Kendall & Beidas).

Despite the pitfalls, numerous advantages make schools a preferred setting for addressing the mental health needs of youth. First and foremost, schools are the most youth-accessible location because this is where children and adolescents spend most of each day (New Freedom Commission on Mental Health, 2003). From an ecological contextual perspective (Bronfenbrenner, 1979) that considers environmental influences of varying degrees, schools are conceptualized as constituting one of the most important proximal influences in a youth's contextual environment. As such, the school setting provides the opportunity to maximize access, affording an increased ability to reach youth by offering interventions "where they are" (Weist, Evans, & Lever, 2003). Thus, when schools provide mental health services they become centers of care that are located within the community. This change can help to eliminate common obstacles that prevent youth from receiving care (Flaherty, Weist, & Warner, 1996), including transportation needs which must often be coordinated around busy and chaotic parental schedules (Storch & Crisp, 2004). Schools also provide a valuable opportunity for multiple methods of evaluation (Owens & Murphy, 2004).

Another crucial advantage to providing services in schools is that they are one of the primary settings in which youth display impairment (Ginsburg, Becker, Kingery, & Nichols, 2008). For youth fraught with anxiety and depression, many of the situations that cause disorder-related interference are interwoven within the school experience. For instance, anxious youth may be apprehensive about separating from parents to attend school, concerned about social interactions among a network of peers, and worried about evaluation on academic performance (McLoone, Hudson, & Rapee, 2006), whereas the school setting may force depressed youth to confront on a daily basis the aspects of life they are depressed about (e.g. absence of meaningful peer relationships or academic underachievement). Thus, school-based interventions are uniquely poised to enhance generalizability by encouraging practice and fostering growth in the very situations that lead to difficulty. Additionally, schools are often comprised of large, diverse populations with a heterogeneous collection of presenting difficulties. All of these factors contribute to schools demonstrating the type of "ecological validity" (Owens & Murphy, 2004) that allows interventions to be investigated in youths' naturalistic environments, with treatment benefits realized in a context that is both clinically and practically meaningful to the everyday lives of children and adolescents.

Providing training to school-based mental health practitioners is another invaluable feature of implementing school-based services, and one that confers a number of benefits (Ginsburg et al., 2008). For instance, their presence in schools allows them to intervene with youth and process problematic situations on a real-time basis. Of particular importance to school systems located in less economically advantaged areas, school-based clinicians can offer programs that are much more affordable as compared to traditional private-practice outpatient or hospital-based services.

Finally, although parental perceptions of the school can sometimes be a source of resistance, they also have the potential to be a positive force. When school-systems are familiar to parents and are viewed as trusted institutions, their reputation may serve to help services seem more acceptable. Further, the naturalistic setting of schools may have the capacity to reduce the stigma that often accompanies mental health treatment in the greater community (Storch & Crisp, 2004). Such a benefit has the potential to translate into important differences regarding access to care, as some research suggests that youth are more likely to utilize school-based services than those that are offered through traditional mental health clinics (Anglin, Naylor, & Kaplan, 1996; Earls, Robins, Stiffman, & Powell, 1989).

Intervention Classification System

A fundamental challenge exists when it comes to comparing and contrasting different school-based interventions, in that the assorted programs differ widely with regard to the intended outcome: Is the goal to prevent the onset of problems, to treat existing problems, or somewhere in between? This variability can make it difficult to draw firm conclusions from the school-based interventions literature when considered as a whole, and leads to an overall lack of clarity between what constitutes prevention and intervention efforts, and how these can be distinguished from one another (Collins & Dozois, 2008). Resolving this confusion and arriving at a consensus on how these terms can be defined in an understandable manner is thus an integral first step toward effective implementation of school-based mental health programs (Kutash, Duchnowski, & Lynn, 2006).

The initial conceptualization of disease prevention was developed approximately a half century ago within a public health framework. The Commission on Chronic Illness (1957) proposed three levels of prevention interventions, which were established with regards to illness

status. *Primary* prevention sought to decrease the incidence of a disorder, or the number of new cases per year. The goal of *secondary* prevention was to lower the prevalence rate of a disorder, or the number of cases already in existence in the population. Finally, *tertiary* prevention was intended to reduce the amount of impairment created by an existing disorder or illness. However, only the primary level seemed to reflect prevention in the true sense of the word, whereas so-called 'prevention' at the secondary and tertiary levels seemed to instead be addressing treatment, with efforts aimed at the alleviation of associated disability and eliminating the disorder entirely.

The insufficiency of this system led to alternative frameworks (Kendall & Norton-Ford, 1982; Gordon, 1983, 1987). Gordon proposed a shift away from the focus on intervening mechanisms, instead suggesting the adoption of a "(cost) risk-benefit" perspective. The resulting model was again a three-tiered system, but one that balanced the projected cost and discomfort of the intervention with the degree of threat posed by an illness. *Universal* preventive measures were those which would be generally acceptable and desirable to all persons in a population due to the benefits greatly outweighing the costs. *Selective* preventive measures, on the other hand, were meant to be appropriate not for everyone, but rather for those who, because of some trait or characteristic (e.g., age, ethnicity, family history, etc.) belong to a subgroup that is at an elevated risk for becoming ill. Finally, *indicated* preventive measures, due to their not being completely benign or minimal in cost, were proposed as being advisable for individual persons, who, based on evaluation, are determined to manifest some abnormality that elevates their risk status.

In developing this system, Gordon (1987) sought to restrict the use of the word "preventive" to "measures, actions, or interventions that are practiced by or on persons who are not, at the time, suffering from any discomfort or disability due to the disease or condition being

prevented" (p. 23) and thereby eliminate the confusion of what had been previously been the domain of tertiary prevention under the 1957 model. Gordon asserted that indicated prevention was differentiated from "treatment" by the latter's aim to be immediately therapeutic for a known condition whereas the former is intended for an asymptomatic individual with a specific risk factor identifier. However, many believed that such a proposed "abnormality" actually reflected the earliest signs of the very illness that the indicated measures were supposed to prevent. As time went on, the two systems were sometimes blended and other times used interchangeably, which only contributed to a greater confusion of terms.

In an effort to finally resolve this issue, a sub-committee of the Institute of Medicine (IOM) – known as the Committee on Prevention of Mental Disorders – was commissioned to develop a new system of classification that would provide a standard for defining interventions. Within this new framework (Mrazek & Haggerty, 1994), the term "prevention" is applied only to those interventions implemented prior to the actual onset of disorder and the term "treatment" refers to interventions that are therapeutic in nature and delivered to people who presently meet diagnostic levels for disorder. Incorporating some of Gordon's (1983, 1997) terminology, the IOM (1994) established the following definitions: "Universal preventive interventions" are programs intended to be administered to the general public or entire population groups, and are not based on any identified risk status; "selective preventive interventions" are meant to be delivered to individuals who have a higher than average risk for developing a mental disorder on the basis of various risk factors – hereditary, psychological, social – already established as associated with its onset; finally, "indicated preventive interventions" seek to help individuals deemed as being high-risk who exhibit measurable symptoms that signal the impending onset of mental disorder, but who do not presently meet sufficient criteria to warrant a clinical diagnosis.

Despite acknowledging that the boundaries proposed by this classification system may be somewhat blurred in the realm of clinical practice, Mrazek and Haggerty (1994) assert that such a system is necessary to enable accurate collection of information and to ensure a common language for communicating about this type of work. Indeed, this system of classification has been widely adopted and is now the most agreed upon method for defining the level of intervention. Despite this step forward, however, school-based interventions continue to experience the type of blurred boundaries predicted by Mrazek and Haggerty. For example, manifesting depressive symptoms could represent a risk factor for a later depressive disorder or may signal the presence of an already existing disorder with unclear presentation. In this way, when the onset of a disorder is unclear, the distinction between indicated prevention and early treatment is much less clear (Albee & Gullotta, 1986).

Another example may be seen in Australia – a country where many school-based interventions have been conducted – where the mental health policy framework has adapted the IOM's (1994) spectrum of interventions and created a new term known as "early intervention" which is defined as including indicated prevention and early treatment (Commonwealth Department of Health and Aged Care, 2000). Thus, early intervention approaches do not exclude individuals who score in the clinical range and meet full disorder status. Though such an approach has its own clinical utility, from a research standpoint it once again melds prevention and treatment together in the very way that the classification efforts described above strove to delineate. Consequently, when the findings of early intervention studies are presented, it can be quite difficult to make conclusions about whether the effects reported reflect a preventive or a disorder-reduction impact, as the blending of participants has the capacity to make these indistinguishable.

Previous Reviews and Meta-Analyses

To what extent, then, have interventions for anxious and depressed youth actually been applied to meet the mental health needs of youth, and what is the status of the collective set of outcomes? The answer to this question requires a synthesis of available information on the topic, though to date, this objective has not yet been fully achieved. The commitment to this type of synthesizing work has increased in recent years, and the collection of reviews of the treatment literature is expanding.

Merry, McDowell, Hetrick, Bir, and Muller (2004) conducted a Cochrane meta-analytic review of randomized controlled trials for the prevention of depression in youth, excluding any who met DSM IV or ICD 10 criteria for depression and/or fell into the clinical range on standardized, validated and reliable rating scales of depression at the start of a study. Results were restricted to psychological interventions only (excluding one educational and two psychoeducational interventions), of which 10 were universal prevention studies and 8 were targeted prevention studies. However, out of 21 studies identified as eligible for inclusion, only 13 were reported to have had data suitable for extraction and inclusion in the meta-analysis. For two studies with an active attention or placebo comparison group, there was no evidence of effectiveness. For studies using no intervention, waitlist, or “usual care” comparison groups, the authors found significant reductions in scores on depression rating scales for targeted, but not universal prevention efforts; however, this was true only immediately after the intervention, with pooled data not supporting effectiveness at longer-term follow-up. Finally, results regarding differential effects by gender were conflicting, showing some advantages for boys over girls at postintervention, but not at any follow-up assessments.

Horowitz and Garber (2006) conducted a meta-analytic review of 30 studies evaluating prevention efforts for depressive symptoms amongst children and adolescents. These authors found that, at post-intervention, selective prevention programs yielded significantly greater effect sizes than those of universal prevention, however no significant differences between indicated prevention programs and either selective prevention or universal prevention were identified. Moderators of outcome were also explored. There was no effect found for the number of sessions included in the program. At post-intervention, it was found that studies with a greater percentage of females had significantly larger effect sizes, though this effect was reduced to nonsignificance when two studies with college-age students were removed. A similar pattern was found in terms of age, whereby significantly greater effects found for older participants were no longer obtained when two studies with college-age students were omitted. Such results suggest that the moderating effects of age and gender in the two studies with college-age participants may not be representative of relationships in outcomes obtained with younger school-age subjects, and that further exploration is needed with this latter population.

In a meta-analysis by Ishikawa, Okajima, Matsuoka, and Sakano (2007), 20 randomized controlled trials of CBT for anxiety disorders in youth were examined to evaluate the degree of therapeutic gain. These authors found an unclear pattern of results regarding the inclusion of parents or family into treatment. They noted that, due to the small number of studies included in their review, more work is needed to identify the influence of CBT for anxiety disorders on comorbid depression symptoms. Results also suggested essentially no difference between treatments employing 10 or fewer sessions as compared to those utilizing 11 or more sessions. Finally, 11 of the 20 studies were conducted in university clinics or hospitals, and, when grouped together, effect sizes for these studies were found to be larger than for the remaining 9 studies

conducted at “other settings.” However in moving toward the expansion of dissemination efforts where efficacious programs are implemented in the community at large, it is exactly these types of *other settings* (e.g., schools) that need to be examined in greater detail.

In some cases, efforts have been made to narrow investigation to only studies of school-based interventions. For instance, Rones and Hoagwood (2003) sought to provide a review of the evidence base for mental health services delivered by schools, examining program evaluations published between 1985 and 1999 that targeted a variety of presenting concerns, and identifying factors associated with success. These included a commitment to consistently high fidelity program implementation, the use of multicomponent programs targeting the ecology of the child, applying multiple modalities to enhance learning of content, integrating the program into the general classroom curriculum, and ensuring developmental appropriateness of the programs. Of note, these authors found there to be a lack of treatment effects for even the most prevalent disorders of childhood and indicated that no studies that targeted particular clinical syndromes were identified. Indeed they noted that, “Surprisingly, we did not find any school-based anxiety prevention or intervention programs that met the criteria for entry into this review” (p.238). Neil and Christensen (2007) also sought to review school-based interventions for anxiety and depression. However, as their focus was specifically on the Australian school environment, the review included only studies examining programs developed or evaluated in that country, and did not apply formal meta-analytic procedures.

Recent years have thus seen a fortuitous growth in the number of efforts to synthesize the treatment outcome literature. However, as noted above, gaps remain insofar as most of the studies tending to be reviewed in meta-analytic examinations are frequently of those conducted in research-based clinics (Weisz & Jensen, 2001) and thus are not sufficiently comprehensive to

offer crucial details about which aspects of school-based interventions most importantly contribute to their success. Indeed, in a recent meta-analytic review of evidence-based psychosocial treatments for anxious youth, Silverman, Pina, and Viswesvaran (2008) stated that while there are existing studies to suggest the feasibility of group-based CBT (GCBT) in school settings, "the efficacy of school-based GCBT warrants further research attention" (p.118).

Two very recent efforts have sought to meet the need of an exclusive focus on school-based intervention. Neil and Christensen (2009) conducted a review of 27 randomized controlled trials for the prevention and early intervention of youth anxiety. Similarly, Calear and Christensen (2010) conducted a review of 42 randomized controlled trials of school-based prevention and early intervention for youth depression. Of note, while the former review determined that significant effects for anxiety interventions did not depend on type of comparison control group or program implementer, the latter review found that interventions using attention-control conditions and those led by teachers were associated with fewer significant outcomes. While effect sizes were reported in these reviews, the authors of both state that a formal meta-analysis was not conducted. As such, to this author's knowledge, no meta-analysis has yet been conducted that is devoted entirely to school-based studies of the prevention and treatment of anxiety and depression in children and adolescents.

The Present Study

The current study reviewed school-based interventions for anxiety and depression in children and adolescents. The rationale for examining both of these disorders is that anxiety and depression share significant similarities in terms of their underlying cognitive and behavioral symptomatology. Additionally, research has demonstrated that anxiety often predates and leads to the development of depression (Brady & Kendall, 1992; Cole, Peeke, Martin, Truglio, &

Seroczynsky, 1998). Thus, given the meaningful relationship that exists between the two, it was believed that joint exploration of both disorders is a worthwhile endeavor and one with the potential to make a valuable contribution to the understanding of how to address the difficulties associated with anxiety and depression within the school setting.

Traditional narrative methods of literature review suffer the weakness of being subjective and influenced by potential biases (Glass, 1976). Statistical approaches to the summarization of a topical area have the advantage of being a less biased alternative. Meta-analysis has been advocated as the preferred method of data synthesis for a number of reasons including its broad-scoped comprehensiveness, the capacity to incorporate data that might not otherwise be included, the ability to identify moderator variables, and more focused attention paid to directionality and magnitude of effects rather than an overreliance on significance testing (Rosenthal & DiMatteo, 2001). The Task Force on Statistical Inference of the American Psychological Association emphasizes "that reporting and interpreting effect sizes in the context of previously reported effects is essential to good research" (Wilkinson and APA Task Force on Statistical Inference, 1999, p. 599). Indeed, as a collection of effect sizes constitutes the primary focus of investigation in meta-analysis, conclusions based on this approach are thus not only more objective, but also more methodologically rigorous (Cooper & Rosenthal, 1980) and are now widely accepted as the most appropriate approach to literature review.

The present project was thus an intervention effectiveness meta-analysis (Durlak & Lipsey, 1991) that explored the magnitude of the effect of school-based interventions for the prevention and treatment of anxious and depressive symptomatology among youth. It focused on addressing the following questions: How effective are school-based interventions in reducing anxious and depressive symptoms amongst school-age youth? What are the relationships

between characteristics of the programs that are associated with more or less positive outcomes? Further, do these interventions yield positive effects on associated outcomes such as self-esteem and comorbid psychopathology symptoms?

Specifically, the present study aimed to explore (1) the overall effect sizes of school-based interventions for reducing anxiety and for depression, as well as (2) the degree of relationship between post-intervention anxiety and depression and child characteristics (e.g. age, gender), intervention characteristics (e.g. number of sessions, program implementers, level of intervention), and study methodology (e.g. control group type, format of outcome measurement, outcome reporter informant).

The primary hypothesis was that combining all studies of school-based interventions for anxiety and depression delivered to children and adolescents would reveal overall statistically significant summary effect sizes, and that these summary effect sizes would be significantly greater than those obtained for youth in control groups. Additional hypotheses included the following: (1) the effect size found for treatment studies would be greater than that found for prevention studies; (2) the effect size found for selective and indicated prevention studies would be greater than that found for universal prevention studies; and (3) increasing duration of intervention would be associated with larger magnitude of effects.

The process and findings of the present meta-analysis are reported below and are guided by the standards developed at the Quality of Reporting of Meta-analyses conference (QUORUM; Moher et al., 1999).

CHAPTER 2

METHOD

For the purposes of this review, an *intervention* was defined as a program of content delivered to students by one or more individuals trained in its implementation with the purpose of either a) preventing the onset or exacerbation of anxiety or depressive symptoms or disorders, or b) alleviating the symptoms and/or severity of already-existing anxious and depressive disorders and associated symptomatology.

Searching

Studies for the present meta-analysis were identified through the use of several search strategies. Firstly, computerized searches were conducted using the PsycINFO and PubMed research databases. The following search terms were utilized to identify potentially appropriate studies: *anxiety, depression, prevention, treatment, intervention, school, school-based, children, adolescents,* and *youth*. The search parameters were limited so as to obtain articles reporting on empirical studies between the 20 year time frame of January 1, 1990 and December 31, 2009. The final online search date was 06/01/2010. Secondly, the reference sections of articles found via these computerized searches as well as those of previously conducted meta-analyses were reviewed to ensure that potentially relevant studies not identified by computerized search were included – this approach was particularly useful for locating evaluations conducted as dissertations that were never published. Third, the tables of contents for journals published between 2005 and 2009 that typically include studies on child psychopathology as it pertains to youth in schools were reviewed to identify other potential studies that may have been missed by the methods of searches described above. These journals included: *School Psychology Review, Journal of School Psychology, School Psychology Quarterly, Journal of Consulting and Clinical*

Psychology, *Journal of the American Academy of Child and Adolescent Psychiatry*, *Archives of General Psychiatry*, *Behaviour Research and Therapy*, *Development and Psychopathology*, *Journal of Abnormal Child Psychology*, *Journal of Abnormal Psychology*, *Journal of Child Psychology and Psychiatry*, *Journal of Clinical Child and Adolescent Psychology*, and *Psychological Bulletin*. Finally, in an effort to reduce the threat of publication bias (see below), several prominent researchers in the field were consulted to ensure inclusion of relevant published and unpublished studies that had not yet been identified.

Selection

Studies were selected as appropriate for inclusion in this review if they met the following criteria. First, the study had to explore the outcome of an intervention for anxiety or depression that was implemented within a school system. Second, the intervention must have been cognitive-behavioral in nature. Third, participants had to be school-age children and adolescents currently enrolled in grades K through 12. Fourth, the size of the samples under investigation had to be sufficiently large enough to allow for the types of statistical analyses required for meta-analytic purposes. As such, single-subject case studies were not included in this review. Fifth, the study must have provided quantitative analyses of outcome data, with the specific statistical information needed for the calculation of effect sizes, or author contact information to obtain the additional data needed. Last, although the content of the intervention may have been delivered in any language, the study must have been written in English. In order to preserve assumptions of independence, studies were excluded if the sample under investigation overlapped either in whole or in part with the sample of participants from another study which met inclusion criteria. When this was suspected, authors were contacted to confirm participant sample overlap. In such instances where this was the case, the study conducted first temporally was chosen for inclusion,

as youth who were included in later examinations were more likely to have entered with prior exposure to psychotherapeutic strategies.

Validity Assessment

The meticulous standards of the peer-review process have led to the suspicion that published studies are more likely to have obtained statistically significant findings as compared to those that never reach the pages of a journal (Sterling, 1959). Studies determined to be lacking in methodological rigor, that failed to reject the null hypothesis, or whose authors decide not to pursue publication for whatever reason contribute to a number of investigations not accounted for by the research literature. It is thus acknowledged that conventional search strategies, such as those described above, may artificially inflate the likelihood of obtaining support for the primary hypothesis of the present study due to overreliance on published – and thus presumably significant – "findings." This issue has become known as the "File Drawer Problem" (Rosenthal, 1979), and refers to the well-supported notion that the studies capable of being retrieved and subsequently making it into a meta-analysis are in reality not likely to be a random sample of all studies actually conducted to explore a particular content area (Rosenthal, 1991). The result of this "file drawer problem" is the emergence of publication bias, whereby smaller studies which have insufficient power to detect an effect or are deemed less important are omitted and not included in the calculation of a mean effect size. Consequently the findings reported in the collection of published studies may not reflect an accurate representation of the results that have truly been obtained. Indeed, despite research suggesting that a majority of meta-analysts and methodologists believe that unpublished material should be included in systematic reviews, only about a third of published meta-analyses were found to include such unpublished work (Cook, Guyatt, & Ryan, 1993).

In the present meta-analysis, the file-drawer problem was addressed in two ways. The first was by not restricting studies to only those evaluations which have been published, as is commonly seen as an inclusion criterion in many meta-analyses. As such, efforts were made to obtain unpublished data from relevant investigations discovered at conference presentations, queries sent to authors of unpublished work cited in article reference lists, via communication with experts in the field, and through a general request for assistance in identifying pertinent unpublished research distributed to the email listserv of the Association for Behavioral and Cognitive Therapies (ABCT).

While these strategies yielded some positive results, the recognition that these efforts could not possibly lead to a comprehensive inclusion of all relevant unpublished studies led to the second approach for addressing the file-drawer problem. Rosenthal (1979) recommends calculating the number of studies averaging null results that never reached publication that would have to exist before the overall probability of a Type I error can be shifted to any desired level of significance, for example $p = .05$. The resultant number of studies is referred to as the "tolerance for future null results," and Rosenthal asserts that the meta-analyst must then evaluate this data to determine whether overall conclusions are threatened by an excessively small tolerance level. If the overall level of significance obtained through the meta-analysis is reduced down to the "*just significant*" level by incorporating a only few additional null results, then results are not deemed to be robust to the file drawer problem.

The way that this is accomplished is through the calculation of a term known as "Fail-Safe N", which reflects the exact number of studies needed to shift a significant overall *p* value beyond a predetermined critical value into the range of nonsignificance (Rosenthal, 1979). Orwin (1983) provided an adaptation to this formula to make the failsafe N applicable to effect sizes.

This allows for the determination of the number of studies with a small magnitude effect size that would be needed to reduce the mean effect size to a specified criterion level. The resulting value offers an estimation of how resistant to null effects the calculated mean effect size is:

$$k_0 = k\left[\frac{\overline{ES}_K}{\overline{ES}_C} - 1\right]$$

In this approach to obtaining the failsafe N (Lipsey & Wilson, 2001), k_0 is the number of studies with an effect size value of zero that would be needed to reduce the weighted mean effect size (represented by $\overline{ES}_K$) to $\overline{ES}_C$, which is the criterion effect size level, and *k* is the number of studies in the meta-analysis that contributed to the calculation of the weighted the mean effect size.

Data Abstraction

All studies were coded by the first author, a doctoral candidate in clinical psychology. Quantitative data from measures examining constructs of interest were entered into a Microsoft Excel (Microsoft Corp., Redmond, WA) database, with algorithms programmed in to calculate effect sizes (ESs). Two graduate students in clinical psychology served as independent coders following training that included a) instruction about the project, b) practice coding, and c) trained-to-criterion testing. This process included an initial 90 minute meeting with the principal investigator (PI) to discuss the school-based intervention literature, important themes driving this field of research (e.g. shifting focus from efficacy to effectiveness), and the categories of interest to be coded for the project. One study was chosen for initial independent practice coding, after which trainees met with the PI for one hour for detailed review of coding selections for each variable and to address questions. Coders were then assigned a group of 4 studies to code independently, and subsequently met with the PI for 90 minutes to review coding discrepancies.

These five studies were drawn due to their representativeness of the sample as a whole. Finally, coders were assigned a randomly selected set of 20 studies – 10 each, from those focusing on anxiety and those focusing on depression – from the overall sample of studies to code independently; the five studies described could not be selected. Inter-rater reliability, calculated using Intra-class correlation (ICC) for variables capturing continuous measures of outcome data and Cohen's kappa (κ) for categorical variables, reached accepted standards for all coded variables included in the current project's analyses (see Results section).

When data necessary for the computation of ESs or related variables of interest was not presented in manuscripts, a concerted effort was made to obtain the needed information directly from the study authors. In some instances, no response was received, authors could not be located for contact (e.g. had retired or changed place of employment), or requested data was no longer available (e.g. for studies conducted during early-to-mid 1990s). Consistent with the study selection criteria described above, if missing quantitative data for effect sizes could not be obtained through communication with authors, that study was excluded.

Measures

A number of self-report symptom inventories were utilized across studies to assess an intervention's effects on youths' experience of anxiety and depression.

Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1978). One of the outcome measures most frequently implemented to assess anxiety in the selected studies, the RCMAS is appropriate for youth between the ages of 6 and 19 years of age, and contains 37 self-report items that assesses chronic or trait anxiety in. All items are presented in a yes-no format and a Lie scale can be calculated from 9 RCMAS items. The other 28 items of the RCMAS reveal three anxiety factors: Physiological Symptoms, Worry and Oversensitivity, and

Social Concern-Concentration. The RCMAS has been found to have good reliability (Reynolds & Paget, 1983) and reasonable construct validity (Reynolds, 1982).

Spence Children's Anxiety Scale (SCAS; Spence, 1998). Another anxiety measure that was used by a considerable number of the school-based anxiety intervention studies to evaluate outcome, the SCAS is a child self-report measure comprised of 38 items which assess for a broad range of anxiety symptoms common amongst youth and consistent with several of the anxiety disorders described in the DSM. Spence (1998) has found this measure to demonstrate strong psychometric properties, with a co-efficient alpha of 0.92, a 6-month test-retest reliability of 0.60, and it correlates well with other measures of anxiety.

Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher et al., 1997). The SCARED is a 41-item measure of a child's anxiety designed for use in children ages 8 and older. Responses to items are given using a 3-point Likert-type scale that describes the degree to which statements are true about the child (e.g. not true or hardly ever true, somewhat true or sometimes true, very true or often true) and it can be completed by either the child or his/her parent. The SCARED demonstrates good internal consistency, test-retest reliability, discriminative validity, and moderate parent-child agreement (Birmaher et al., 1999).

Social Phobia and Anxiety Inventory for Children (SPAI-C; Beidel, Turner, & Morris, 1995). The SPAI-C uses 26 items to assess the specific somatic symptoms, cognitions, and behavior that commonly occurs in conjunction with social anxiety in children and adolescents across a range of potentially fear-producing situations. Questions are answered on a 3-point Likert-type scale ranging from 0 (never) to 2 (most of the time or always). The SPAI-C has been shown to be internally consistent (Cronbach's $\alpha = .95$) and have test-retest reliabilities of 0.86 and 0.63 at 2-week and 10-month intervals (Beidel et al., 1995). The SPAI-C has correlated

significantly with the Social Anxiety Scale for Children-Revised, the number of socially distressing events as reported by a daily diary, independent observer ratings of behavioral skill, and parental reports of children's social fears (Beidel, Turner, & Fink, 1996; Beidel, Turner, Hamlin, & Morris, 2000; Morris & Masia, 1998). This measure also has been found to discriminate children with social phobia from those with other anxiety disorders, with externalizing disorders, and with no psychiatric disorders.

Multidimensional Anxiety Scale for Children (MASC; March, 1997). The MASC is comprised of 39 items with a four-point response format. Items load into four subscales, or factors, including physical symptoms, social anxiety, harm avoidance and separation anxiety. The MASC has demonstrated good psychometric properties, including interrater concordance as well as convergent and divergent validity (March, Parker, Sullivan, Stallings, & Conners, 1997).

State-Trait Anxiety Inventory (STAI, Spielberger, 1970). The STAI is an instrument with 20 items assessing state anxiety and 20 items used to assess trait anxiety. The STAI thus provides measures both of how much anxiety an individual is feeling at the time of assessment, as well as how much the individual generally experiences on a regular basis. It has shown to have good reliability and validity (Bedell & Roitzch, 1976; Chaplin, 1984). A parallel version, the *State-Trait Anxiety Inventory for Children* (Spielberger, 1973) is also available for use with youth in earlier childhood.

Social Anxiety Scale for Children (SAS-K; Dekking, 1983). The SAS-K was used to measure intervention outcome in one study (Bokhorst, Goossens, & deRuyter, 1995), which indicated selection of this measure due to the availability of national norms and good psychometric properties (Dekking, 1983). Designed for use with children between nine and

twelve years of age, the SAS-K consists of 46 items assessing social anxiety, with higher scores reflecting greater impairment.

Social Anxiety Scale for Adolescents (SAS-A; La Greca, 1998). The SAS-A was used to measure change produced by intervention in one study (Ginsburg & Drake, 2002). It was described as a modified version of the *Social Anxiety Scale for Children-Revised* (La Greca & Stone, 1993), designed to be appropriate for use with youth ages 12-18. The SAS-A contains 22 items, answered on a five-point Likert-style format indicating the degree and frequency to which respondents feel that an item describes their feelings. La Greca (1998) has found acceptable psychometric support for the subscales of the SAS-A.

Children's Depression Inventory (CDI; Kovacs, 1985, 1992). The CDI was the tool most widely used by studies for the measurement of depression, and contains 27 items assessing cognitive, affective, and behavioral depressive symptoms. It has been shown to have good internal consistency, moderate retest reliability, and correlates with measures of related constructs (Kovacs, 1985). Psychometric and normative data are available (Finch, Saylor, & Edwards, 1985; Finch, Saylor, Edwards, & McIntosh, 1987). Scores have been found to predict depressive disorders and differentiate depressive disorders from other disorders (Timbremont, Braet, & Dreessen, 2004). Of note, however, in some studies (e.g., Kowalenko et al 2005; Shochet et al., 2001), permission was refused for the item relating to suicidal ideation to be presented to participating students, so it was omitted, leaving 26 items. In these cases, omitting one item, but using Kovacs' (1992) original cut-offs, resulted in a conservative measure of depressive symptomatology.

Studies exploring interventions with older adolescents often opted to use the *Beck Depression Inventory* (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The BDI is a 21-

item scale with a three-point response format, with higher scores signifying greater depression severity. Its reliability and validity have been documented (Stehouwer, 1985).

Center for Epidemiologic Studies-Depression (CES-D; Radloff, 1977) was also used for evaluating youth depression in school-based interventions. The CES-D is a 20-item self-report measure of depressive symptomatology, with responses ranging from 0 (none of the time or rarely) to 3 (most of the time). It has been found to have adequate reliability and validity (Myers & Winters, 2002). Scores range from 0 to 60, with greater scores reflecting higher levels of depression, and a cutoff score of 16 is frequently used to identify clinically significant depressive symptoms.

Reynolds Adolescent Depression Scale (RADS; Reynolds, 1986). The RADS is a questionnaire consisting of 30 items, each answered on a four-point response format, with higher scores reflecting greater depression-related impairment. The RADS has been found to demonstrate good reliability (Dalley, Bolocofsky, Alcorn, & Baker, 1992) and to be highly correlated with other well-validated self-report measures of depression (Reynolds, 1990). A parallel version of this measure, the Reynolds Child Depression Scale (RCDS; Reynolds, 1989) is also available and was used in one study (Jaycox, Reivich, Gillham, & Seligman, 1994) for examining depression in earlier childhood. The RCDS measures the frequency of depression symptoms rather than their severity, and has been found to demonstrate good reliability and correlate well with other valid measures of depression (Reynolds, 1992).

Self-Report Questionnaire—Depression (SBB-DES). The SBB-DES was used in one study (e.g. Pössel, Seemann, & Hautzinger, 2008) and measures depressive symptomatology amongst youth using 26 items answered on a four-point Likert-type scale. The overall score is calculated by obtaining the mean of all items, with higher scores reflecting greater endorsement

of symptoms and thus more severity. Pössel and colleagues (2008) found the SBB-DES to have good internal consistency (Cronbach's α = .92) and adequate test-retest reliability.

Study Characteristics

In addition to coding ES data, information related to research design (e.g. level of intervention, participant demographics, study year) and intervention delivery (e.g. number of sessions, session length, program implementer) was also coded – either as continuous variables or as artificially created categorical dummy variables – so as to determine if ES's fluctuated as a function of any of these variables, thus identifying potential mediators or moderators.

Quantitative Data Synthesis

It was often the case that individual studies utilized and reported on more than one measure for a particular construct (e.g. anxiety, depression, self-esteem). However, treating multiple measures of a unitary construct as distinct entities would violate assumptions of independence that underlie the statistical process of meta-analysis (Rosenthal, 1984). Following the recommendations of Lipsey and Wilson (2000) and consistent with the approach taken by authors of recently published high quality meta-analyses (e.g. Stewart & Chambless, 2009), in order to address this issue, multiple effect sizes for a particular construct within individual studies were averaged. This was done prior to synthesis with effect sizes from other studies so as to ensure that each study would only contribute one single effect size per construct.

Effect Size. For the present meta-analysis, it was necessary to decide between two different forms of effect size to evaluate the effectiveness of interventions. One option that is often used in the literature is a "between groups" approach that is constituted by the differences in scores at a particular point in time among participants receiving an active intervention versus those in a control condition, an approach referred to as the standardized mean difference (SMD;

Lipsey & Wilson, 2001). The use of SMD is sometimes suggested as advantageous because it's incorporation of a control group is said to increase confidence that observed effects are actually attributable to the implemented active intervention, and not just due to such nonspecific epiphenomenal causes such as simply being in a treatment study, receiving therapeutic attention, the passage of time, or the notion that one is working to resolve life problems. Thus, by eliminating these types of possible confounds, between-subjects effect sizes are said to be statistically more robust.

Another option for effect size is the standardized mean gain (SMG; Lipsey & Wilson, 2001), which can be conceptualized as a within-groups measure of effect size, and is used to explore changes in continuous measures of constructs of interest over time (e.g. baseline to posttreatment). As such, the SMG does indeed allow for taking control groups into account, and in a manner that the author believes to more powerful: After calculating the SMG for all intervention groups as well as all control groups across the collection of included studies, the two summary effect sizes can be compared so as to determine if one is statistically significantly greater than the other. In this way, findings are not limited to knowing only if intervention groups have lower scores than control groups at a singular posttreatment timepoint – an effect that could potentially be affected by nonequivalent group scores at baseline – as with SMD. Rather, the use of SMG allows for the determination of whether an intervention group truly yields statistically significantly greater ***change*** in scores over time as compared to a control group, which takes baseline scores into account.

Additionally, because the SMD can only be calculated for studies using control groups, reliance on this measure of effect size would require the exclusion of studies lacking control groups. While such studies are acknowledged to be methodologically less rigorous, to exclude

them would lead to a loss of available data while a main goal of the present review was to maximize comprehensiveness.

In order to calculate the SMG, the correlation between baseline and posttreatment scores is needed; unfortunately, however, most of the studies selected did not provide this value or the raw data necessary to calculate it. As such, following the recommendation offered by Rosenthal (1993) and used in recently published meta-analyses (e.g. Hoffman et al., 2010), a conservative estimate of r=0.7 was utilized.

All calculated effect sizes are represented in the form of Hedges *g*, as the distribution of Cohen's *d* effect sizes may be upwardly biased (Lipsey & Wilson, 2000) if based on a number of studies with small sample sizes (e.g. N < 20). Hedges *g* is obtained by multiplying Cohen's *d* by an adjustment factor:

$$\text{Hedges } g = (d) \text{ X } (1 - \frac{3}{4df - 1})$$

and the magnitude may be interpreted according to the convention established by Cohen (1988), in which effect sizes are considered small (0.2), medium (0.5), and large (0.8). For most of the measures of anxiety and depression utilized, higher scores indicate greater levels of symptomatology. SMG effect sizes were therefore calculated by subtracting post-intervention scores from baseline scores, such that positive values would reflect effects in the expected direction (e.g. reduced over the course of the intervention). Correspondingly, a negative SMG effect size indicates that symptomatology worsened over time.

From the distribution of Hedge's *g* values produced, a summary effect was produced by pooling across studies and calculating an average effect size statistic. Analyses were completed following the procedures outlined by Borenstein, Hedges, Higgins, and Rothstein (2009), developers of the software program Comprehensive Meta-Analysis, which has been utilized in

other recent meta-analytical investigations (e.g. Brunwasser, Gillham, & Kim, 2009; Hoffman, Sawyer, Witt, & Oh, 2010). Additional analyses were conducted using the SPSS Version 18.0 for Windows statistical software package (SPSS, Inc., Chicago, IL).

Testing for Heterogeneity of Effect Sizes. After the combined mean effect size has been obtained, an important question remains: Is the distribution of effect sizes homogeneous (Hedges, 1982)? In other words, do the effect sizes in the set used to obtain the mean effect size all estimate the same population effect size? In a homogeneous distribution, each individual effect size would be expected to differ from the population mean effect size only as a function of sampling error. However, when the null hypothesis of homogeneity is rejected, this indicates that the variability among effect sizes in a distribution is greater than would be expected from sampling error alone (Lipsey & Wilson, 2000). In this case, it must be assumed that the variation in effect sizes is due to a source other than only participant-based sampling error. Such an approach operates from the position that the excess variability in effect sizes is due to random differences that cannot be identified among the distribution of studies. The equation used to conduct the homogeneity test is:

$$Q = \sum_{i=1}^{k} W_i Y_i^2 - \frac{(\sum_{i=1}^{k} W_i Y_i)^2}{\sum_{i=1}^{k} W_i}$$

It is understood that the primary goal of meta-analysis is the synthesis and summarization of existing data for the express purpose of being able to make inferences from the collection of studies being examined. However, two statistical models – fixed effects and random effects – have been developed that differ in their approach to describing the specific universe to which conclusions can be applied. Within this framework, the “universe” refers to the hypothetical assortment of studies that, in principle, could possibly be conducted, and about which we wish to

generalize (Cooper & Hedges, 1994). The fixed-effects model is understood to be a conditional model, because it assumes as fixed the various features of studies that could be related to the effect size parameter. In this way, the universe to which generalizations are made from a fixed effects model approach consists of studies that differ from those analyzed only as a function of having different samples of participants. On the other hand, random effects models do not make this assumption. By not holding fixed (unconditional) the characteristics of studies, generalizations can be made to a universe that contains studies that are not identical to those which were included in the sample of studies analyzed.

It is important to consider the main advantages and drawbacks to each of the possible models. The fixed effects model is commonly utilized for quantitative research synthesis (Cooper & Hedges, 1994) and is described by the National Research Council (1992) as "the rule rather than the exception" since it provides a more powerful test. However, a major limitation is found in trying to determine if generalizations can be made to studies not exactly the same as those included in the meta-analysis conducted. In this case, the analyst is forced to question, "How similar is similar enough?" to allow such generalization (Hedges & Vevea, 1998), which opens the door to subjective interpretations that may lack reliability. Further, fixed effects models may result in the construction of incorrectly narrow confidence intervals (Hunter & Schmidt, 2004; Lipsey & Wilson, 2000). By using a random effects model, inferences are not restricted only to cases with predictor variables already represented in the sample (Cooper & Hedges, 1994) and comparisons can be generalized to studies that are not identical to those in the sample (Rosenthal & DiMatteo, 2001).

Due to the differences between models described, it is recommended that the choice of model should ultimately depend on the type of inferences the analyst wishes to make (Hedges &

Vevea, 1998). Given the underlying theory, Borenstein and colleagues (2009) suggest that a fixed effects model should only be used if it is believed that all studies are functionally identical and the goal is not to generalize to other populations. As the studies being examined were not functionally identical and given that generalization to studies beyond those under investigation is a principal goal of this project, neither of these tenets apply and so a random effects model approach was adopted (Moses, Mostellar, & Buehler, 2002).

CHAPTER 3

RESULTS

Study Characteristics

A total of 63 studies were included in the present meta-analysis, and were conducted in 11 countries, with the majority taking place in Australia and the United States. A total of 15,211 youth were involved; as some studies jointly targeted anxiety and depression, there were 7,885 youth in anxiety intervention studies and 9,727 involved in studies of interventions for depression. The majority of studies (n = 44) used some form of random allocation to condition at either the level of individual student, class, or entire school. Fifteen studies conducted assessments only at baseline and post-intervention whereas the remainder evaluated outcomes across a range of follow-up periods ranging from as little as 1 month to up to 4 years later. In some instances, evaluations of long-term outcome that were later published separately from the initial article were used to obtain follow-up data. In terms of intervention classification level, 31 studies were universal prevention, 8 were selective prevention, 6 were indicated prevention, 5 were targeted prevention, 3 were early intervention, and 11 were treatment studies. Of the studies included, 57 were published in peer-reviewed journals and 6 were unpublished manuscripts that were most often doctoral dissertations. The characteristics of each study can be found in Table 1 and Table 2.

Effect Size Distribution

Homogeneity Analysis

The distributions of effect sizes for the collection of studies are represented as stem-and-leaf plots in Tables 3-6. These were evaluated to determine if existing variation can be entirely explained by random sampling error within studies, or whether it reflects true and meaningful

differences between studies. Statistical tests utilizing the *Q*-statistic (Borenstein et al., 2009) and I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003) were conducted. The *Q*-statistic evaluates the null hypothesis that all studies share a common effect size, whereas the I^2 statistic estimates the proportion of observed variance that reflects real differences in effect size. With regards to the latter, 25%, 50%, and 75% are suggested standards against which to compare an obtained I^2 statistic, reflecting "low," "moderate," and "high" amounts, respectively, of how much variance is accounted for by real differences.

For the collection of studies exploring anxiety interventions, the null hypothesis that all studies share a common effect size was rejected and it was concluded that the true effects vary ($Q = 228.73$, $p < .0001$). Further, the I^2 statistic indicates that nearly 90% of the observed variance is accounted for by real differences. For the collection of depression intervention studies, the null hypothesis was similarly rejected, concluding that there is sufficient evidence that true effects vary ($Q = 128.11$, $p < .001$). In this case, the I^2 statistic indicates that 75% of the observed variance in effect sizes is accounted for by real differences.

The homogeneity results support the a priori decision to conduct the meta-analysis according to a random effects model. These statistics show that the studies do not share one common (true) effect size, and that the factors that could influence effect size are, indeed, not the same in all the studies included. These findings direct us to explore what the reasons might be for the variance, with such efforts described below using subgroup analysis.

Coder Agreement

The coders achieved a high level of inter-rater reliability. Agreement was particularly high (ICC > .90) for continuous ES outcome data, and was in the acceptable range ($.70 > \kappa \geq 1.0$) for all categorical variables of interest.

Quantitative Data Synthesis

Standardized Mean Gain Effect Size. For the 27 studies evaluating school-based interventions for anxiety that had baseline and posttreatment data, the summary pre-post effect size estimate (Hedge's *g*) for those receiving the intervention was 0.50 (95% CI [0.40, 0.60], $p < .001$) for reducing anxious symptomatology (Table 1). Of those studies, 22 implemented control groups, for which the summary pre-post effect size estimate was 0.19 (95% CI [.09, .30], $p < .001$) in terms of a decrease in anxiety symptoms over time (Table 1). A comparison of these summary effect size estimates illustrates a significant difference between the two ($Z = 4.16$, $p < .001$), with intervention group participants demonstrating greater reductions in anxious symptomatology from baseline-to-postintervention than controls (see Figure 1).

For the 39 studies evaluating school-based interventions for depression that had baseline and posttreatment data, the summary pre-post effect size estimate for those receiving an intervention was 0.30 (95% CI [0.22, 0.27], $p < .001$) for reducing symptoms of depression (Table 2). Of those studies, 35 included control groups, for which the summary pre-post effect size was 0.09 (95% CI [0.03, 0.15], $p < .05$) for decreases in depressive symptoms over time (Table 2). A comparison of these summary effect size estimates illustrates a significant difference between the two ($Z = 4.18$, $p < .001$), with intervention group participants demonstrating greater reductions in depressive symptomatology from pre-to-post than controls (see Figure 2).

For the calculation of a fail-safe *N*, a criterion effect size of 0.10 was selected as the level at which results would no longer be considered meaningful, as this represents half of what Cohen's standards for effect size interpretation suggest are small effect sizes (Cohen, 1988). Using this value, results indicate that 108 anxiety intervention studies with an effect size of zero

would have to remain unidentified (“in the file drawer”) to reduce the summary effect size for anxiety interventions from 0.50 to 0.10. With regards to depression, 77 conducted and un-obtained studies with an effect size of zero would have to exist to reduce the mean effect size for depression interventions from 0.30 to 0.10. Although the criterion effect size of 0.10 is acknowledged to be quite small, from the public health perspective such effects can be meaningful, as moving the distribution of symptoms in the population by even a small amount will often correspond to a reduction in the number of overall cases of disorder (Andrews, Szabo, & Burns, 2002). These fail-safe *N* findings therefore suggest that the obtained summary effect sizes are relatively robust, would not be considerably altered by presence of a few unidentified studies reaching null effects, and thus are an fairly accurate representation of the research base.

Outcomes Over Time

Anxiety Interventions.

3-Month Follow-Up. There were five anxiety intervention studies that assessed outcomes three months after the end of the intervention. For these trials, there was a significant mean SMG effect size from baseline to 3-month follow-up for those who had received the intervention of 0.67 (95% CI [0.43, 0.91], $p < .001$). One of these 5 trials did not make use of a control group, and in another two, the control group received the intervention after the main study period. This left two studies with baseline and 3-month follow-up data for youth in control groups, for which there was a nonsignificant mean SMG effect size of 0.09 (95% CI [-0.10, 0.29], $p > .05$). Direct comparison demonstrates that youth receiving interventions experienced significantly greater reductions ($Z=0.68$, $p < .001$) in anxious symptomatology from baseline to 3-month follow-up than youth in control groups (see Figure 1).

6-Month Follow-Up. There were seven anxiety intervention studies that assessed outcomes six months after the end of the intervention. For these trials, there was a significant mean SMG effect size from baseline to 6-month follow-up for those who had received the intervention of 0.57 (95% CI [0.39, 0.75], $p < .001$). In one of these seven trials the control group received the intervention after the 3-month follow-up period. This left six studies with baseline and 6-month follow-up data for youth in control groups, for which there was a significant mean SMG effect size of 0.44 (95% CI [0.15, 0.72], $p < .01$). Direct comparison, however, indicated that there was no significant difference ($Z=0.77$, $p > .05$) in anxiety reduction from baseline to 6-month follow-up between youth receiving active interventions and those assigned to control groups (see Figure 1).

12-Month Follow-Up. There were four anxiety intervention studies that assessed outcomes twelve months after the end of the intervention. For these trials, there was a significant mean SMG effect size from baseline to 12-month follow-up for those who had received the intervention of 0.68 (95% CI [0.47, 0.89], $p < .001$). Each of these trials had control groups that were followed over the 12-month follow-up period, at which point there was a significant mean SMG effect size of 0.56 (95% CI [0.30, 0.81], $p < .001$). Direct comparison indicated that there was no significant difference ($Z=0.73$, $p > .05$) in anxiety reduction from baseline to 12-month follow-up between youth receiving active interventions and those assigned to control groups (see Figure 1).

***Depression Interventions*.**

3-Month Follow-Up. There were 13 depression intervention studies that assessed outcomes three months after the end of the intervention. For these trials, there was a significant mean SMG effect size from baseline to 3-month follow-up for those who had received the

intervention of 0.44 (95% CI [0.27, 0.61], $p < .001$). In three of these trials the control group received the intervention after the main study period. This left 10 studies with baseline and 3-month follow-up data for youth in control groups, for which there was a significant mean SMG effect size of 0.24 (95% CI [0.07, 0.42], $p < .05$). Direct comparison revealed that there were no significant differences ($Z=1.61$, $p > .05$) in reduction of depressive symptomatology by 3-month follow-up for youth receiving active interventions compared to controls (see Figure 2).

6-Month Follow-Up. There were 20 depression intervention studies that assessed outcomes six months after the end of the intervention. For these trials, there was a significant mean SMG effect size from baseline to 6-month follow-up for those who had received the intervention of 0.31 (95% CI [0.21, 0.40], $p < .001$). One of these studies did not employ a control group and in another study the control group received the intervention after the main study period. This left 18 studies with baseline and 6-month follow-up data for youth in control groups, for which there was a significant mean SMG effect size of 0.12 (95% CI [0.03, 0.20], $p < .01$). Direct comparison revealed that youth receiving active interventions experienced significantly greater reductions ($Z= 2.96$, $p < .01$) in depressive symptomatology from baseline to 6-month follow-up compared to controls (see Figure 2).

12-Month Follow-Up. There were seven depression intervention studies that assessed outcomes one year after the end of the intervention. For these trials, there was a significant mean SMG effect size from baseline to 12-month follow-up for those who had been assigned to intervention groups of 0.36 (95% CI [0.14, 0.57], $p < .01$). Each of these trials had control groups which were assessed at the 12-month follow-up period, at which point there was a significant mean SMG effect size of 0.27 (95% CI [0.02, 0.51], $p < .05$). Though youth in both the intervention groups as well as the control groups demonstrated a significant SMG effect size,

direct comparison revealed that the former did not experience significantly greater reductions (Z= 0.55, p > .05) in depressive symptomatology than the latter from baseline to 12-month follow-up (see Figure 2).

Intervention Classification Level.

Anxiety Interventions.

Universal Interventions. Twelve of the 14 studies evaluating universal school-based interventions for anxiety had baseline and postintervention data. For these trials, the summary pre-post SMG effect size estimate (Hedge's *g*) for those receiving the intervention was 0.32 (95% CI [0.22, 0.43], p < .001) for reducing anxious symptomatology. Eleven of those 14 studies had control groups with baseline and postintervention data, for which the summary pre-post SMG effect size estimate on continuous measures of anxiety was 0.19 (95% CI [0.04, 0.33], p < .05). Direct comparison highlights that both youth receiving interventions and controls experienced significant reductions in anxious symptomatology from baseline to postintervention; as such, there was no significant difference between the two groups (Figure 1) in terms of the amount of change (Z=1.54, p > .05).

Selective Interventions. All (n=4) of the studies evaluating selective school-based interventions for anxiety had baseline and postintervention data. For these trials, the summary pre-post SMG effect size estimate for those receiving the intervention was 0.53 (95% CI [0.32, 0.74], p < .001) for reducing anxious symptomatology. All four studies had control groups with baseline and postintervention data, for which the summary pre-post SMG effect size estimate on continuous measures of anxiety was 0.04 (95% CI [-0.06, 0.14], p > .05). Direct comparison highlights that youth receiving selective interventions experienced significantly greater reductions in anxious symptomatology (Figure 1) than did controls (Z=4.09, p < .001).

Interventions for Youth with Elevated Symptom Levels. There were four studies evaluating school-based interventions for youth with elevated (though subclinical) levels of anxiety (e.g. indicated & targeted approaches) that had baseline and postintervention data. For these trials, the summary pre-post SMG effect size estimate for participants receiving the intervention was 0.79 (95% CI [0.44, 1.15], $p < .001$) for reducing anxious symptomatology. Only two of these studies had control groups with available baseline and postintervention data, for which the summary pre-post SMG effect size estimate on continuous measures of anxiety was -0.02 (95% CI [-0.24, 0.21], $p > .05$). Direct comparison highlights that youth with elevated levels of anxiety receiving active interventions experienced significantly greater reductions (Figure 1) in symptomatology than did controls ($Z=3.77$, $p < .001$).

Treatment Interventions. Baseline and postintervention data were available for all intervention group youth in the five studies evaluating school-based treatment for youth meeting diagnostic criteria for an anxiety disorder. For these trials, the summary pre-post SMG effect size estimate for those receiving the intervention was 1.10 (95% CI [0.71, 1.50], $p < .001$) for reducing anxious symptomatology. Three of these studies had control groups with available baseline and postintervention data, for which the summary pre-post SMG effect size estimate on continuous measures of anxiety was 0.35 (95% CI [0.09, 0.62], $p < .01$). Youth receiving treatment interventions demonstrated significantly greater reductions in symptomatology (Figure 1) than did controls ($Z=3.11$, $p < .01$).

Depression Interventions.

Universal Interventions. There were 21 studies evaluating universal school-based interventions for depression that had baseline and postintervention data. These trials yielded a significant summary pre-post SMG effect size estimate for those receiving the intervention of

0.16 (95% CI [0.09, 0.22], $p < .001$) for reducing depressive symptomatology. All of these studies had control groups with baseline and postintervention data, which yielded a nonsignificant summary pre-post SMG effect size estimate on continuous measures of depression of 0.03 (95% CI [-0.03, 0.09], $p > .05$). Direct comparison demonstrated that, while baseline to postintervention effects for youth receiving universal interventions were small, they were significantly (Figure 2) greater than controls ($Z=2.77$, $p < .01$).

Selective Interventions. All (n=4) of the studies evaluating selective school-based prevention efforts for depression had baseline and postintervention data for both intervention and control groups. For these trials, the summary pre-post SMG effect size estimate for those receiving the intervention was 0.38 (95% CI [0.23, 0.54], $p < .001$) for reducing anxious symptomatology. Control groups studies yielded a summary pre-post SMG effect size estimate for continuous measures of depression of 0.11 (95% CI [-0.11, 0.32], $p > .05$). Direct comparison revealed significant differences in depressive symptomatology between selective interventions versus controls ($Z=2.08$, $p < .05$).

Interventions for Elevated Symptom Levels. There were seven studies evaluating school-based interventions for youth with elevated levels of depression (e.g. indicated & targeted approaches). For these trials, the summary pre-post SMG effect size estimate for those receiving the intervention was 0.46 (95% CI [0.23, 0.69], $p < .001$) for reducing depressive symptomatology. One of these studies did not have a control group, and thus there were six studies with available baseline and postintervention data for controls. The summary pre-post SMG effect size estimate on continuous measures of depression for these controls was 0.26 (95% CI [0.04, 0.48], $p < .05$). Direct comparison revealed that youth with elevated levels of

depression at baseline receiving active interventions did not experience significantly greater reductions (Figure 2) in symptomatology by postintervention than did controls (Z=1.19, p > .05).

Treatment Interventions. Baseline and postintervention data were available for all intervention group youth in the five studies evaluating school-based treatment for youth meeting diagnostic criteria for a depressive disorder. For these trials, the summary pre-post SMG effect size estimate for those receiving the intervention was 1.06 (95% CI [0.41, 1.71], p < .01) for reducing depressive symptomatology. Four of those five studies had a control group with available baseline and postintervention data, for which the summary pre-post SMG effect size estimate on continuous measures of depression was 0.26 (95% CI [0.02, 0.49], p < .05). Youth receiving treatment interventions demonstrated significantly greater reductions in depressive symptomatology (Figure 2) than did controls (Z=2.29, p < .05).

Secondary Analyses

Anxiety Interventions' Effect on Depression. From among the studies with interventions intending to target anxiety, there were seven trials which also explored the program's effects on depressive symptomatology. For these studies, there was a significant mean SMG effect size from baseline to postintervention for those who had received the intervention of 0.24 (95% CI [0.06, 0.42], p < .01). One of these studies did not employ a control group, and another did not present data for the control group at postintervention. This left five studies with baseline and postintervention data for youth in control groups, for which there was a significant mean SMG effect size of 0.20 (95% CI [0.11, 0.28], p < .001). Direct comparison revealed that, although youth in anxiety intervention groups experienced significant reductions in depressive symptoms, they did not demonstrate significantly greater reductions in depressive symptoms than controls (Z= 0.46, p > .05).

Depression Interventions' Effect on Anxiety. From among the studies with interventions intended to target depression, there were nine trials which also explored the program's effects on anxious symptomatology. For these studies, there was a significant mean SMG effect size from baseline to postintervention for those who had received the intervention of 0.44 (95% CI [0.24, 0.64], $p < .001$). Two of these studies did not employ a control group, leaving seven studies with baseline and postintervention data for youth in control groups, for which there was a significant mean SMG effect size of 0.15 (95% CI [0.01, 0.29], $p < .05$). Direct comparison revealed that depression intervention groups experienced significant greater reductions in anxious symptoms, than did those youth who were assigned to control groups ($Z = 2.38$, $p < .05$).

Intervention Effects on Related Constructs. This meta-analysis sought to determine whether interventions focused on improving symptomatology for a specific disorder also led to improvements in related constructs. Anxiety intervention studies seldom assessed additional constructs other than depression (as reported above), and so the analyses below focus only on related constructs amongst depression intervention studies.

Self-Esteem. There were nine depression intervention studies that assessed the impact of a particular program on youth self-esteem. For these trials, there was a nonsignificant mean SMG effect size from baseline to postintervention for those who had received the intervention of 0.13 (95% CI [-0.12, 0.38], $p > .05$). All nine of these studies used a control group, for which there was a nonsignificant mean SMG effect size of -0.07 (95% CI [-0.31, 0.17], $p > .05$). Direct comparison revealed that youth receiving active interventions did not experience significantly greater improvements in self-esteem than controls ($Z = 1.12$, $p > .05$).

Hopelessness. Six of the included studies examining interventions for depression incorporated measures of hopelessness and provided data at baseline and postintervention. The

mean pre-post effect size estimate for those receiving the intervention was significant at 0.20 (95% CI [0.08, 0.32], $p < .01$) for improving hopelessness (e.g. becoming less hopeless about the future). All six of these studies implemented control groups, for which the summary pre-post effect size estimate was 0.16 (95% CI [-0.06, 0.37], $p > .05$). A comparison of these summary effect size estimates ($Z = 0.35$, $p > .05$) indicates that, although intervention group participants demonstrated significant improvements in hopelessness from baseline to postintervention and controls did not, the former did not demonstrate significantly greater changes in hopelessness over time as compared to the latter.

Attributional Style. Thirteen of the included studies examining interventions for depression incorporated measures of attributional style (e.g. explanatory style) and provided data at baseline and postintervention. The mean pre-post effect size estimate for those receiving the intervention was 0.19 (95% CI [0.07, 0.31], $p < .01$) for improving explanatory style (e.g. either making it more optimistic or less pessimistic). Of those studies, all 13 implemented control groups, for which the summary pre-post effect size estimate was 0.05 (95% CI [-0.05, 0.15], $p > .05$). A comparison of these summary effect size estimates ($Z = 1.82$, $p > .05$) indicates that intervention group participants did not demonstrate significantly greater improvements in explanatory style from baseline to-postintervention than did controls.

Moderator Analyses. Are there variables that impact the direction and/or strength of the relationship between treatment status and primary outcomes of anxiety and depression symptomatology? In meta-analysis, examining the significance of differences in effect size between categories of variable is a test of whether that variable is a moderator (Shadish & Sweeney, 1991).

Intervention Implementer. The comparison of mean effect sizes produced when interventions were implemented by school staff versus research staff was limited to evaluation of universal prevention studies for the following two reasons: First, studies evaluating universal prevention interventions were more likely than other intervention levels to have school staff implement a protocol, thus creating a larger sample size of studies. Second, including indicated prevention and treatment studies where research staff were implementers would artificially inflate the summary effect size produced for this group, as these intervention levels are generally associated with larger effect sizes (see above).

Amongst the universal anxiety prevention evaluations, there were six that were led by school staff, who were classroom teachers in five studies and school nurses in one. There were six universal preventive interventions led by research staff, who consisted of psychologists and graduate students in training. Youth receiving universal prevention programs implemented by teachers had a significant SMG effect size of 0.33 (95% CI [0.17, 0.48], $p < .001$), and those in universal prevention programs led by research staff had a significant SMG of 0.41 (95% CI [0.26, 0.55], $p < .001$). A direct comparison of universal prevention interventions implemented by teachers and those led by research staff revealed no significant differences in mean effect size (Z=0.77, $p > .05$).

Amongst the universal depression prevention evaluations, there were six that were led by school staff, and in all of these studies the implementers were teachers. There were eight universal preventive interventions led by research staff, who consisted of psychologists and graduate students in training. Youth receiving universal prevention programs implemented by teachers had a significant mean SMG effect size of 0.18 (95% CI [0.08, 0.28], $p < .01$), and those in universal prevention programs led by research staff yielded a significant SMG of 0.18 (95%

CI [0.00, 0.35], $p < .05$). A direct comparison between universal prevention interventions implemented by teachers and those led by research staff revealed no significant differences in mean effect size (Z=-0.01, $p > .05$).

Intervention Dose. Interventions varied both in terms of the number and duration of sessions included in a program. Intervention "dose" was thus defined as the total number of minutes involved in the intervention, (i.e., number of sessions multiplied by session duration). For the same reasons described earlier, the comparison of mean effect sizes produced by differing levels of intervention dose was limited to evaluation of universal prevention studies. For anxiety and depression, respectively, studies of universal interventions were ordered according to intervention dose. The median dose was used to divide the studies into "low dose" and "high dose" interventions, which were then directly compared to determine if dosage was a moderator of intervention effects.

Twelve studies of universal interventions for anxiety had available baseline and postintervention data. The average low dose was 354.87 minutes and the average high dose of intervention was 682.50 minutes. Youth receiving low dose universal interventions had a significant mean SMG effect size of 0.32 (95% CI [0.14, 0.49], $p < .001$), and those receiving high dose universal interventions had a significant SMG of 0.32 (95% CI [0.19, 0.46], $p < .001$). Direct comparison between low dose and high dose interventions, however, revealed no significant differences in mean effect size ($Z = 0.08$, $p > .05$).

There were 20 studies of universal interventions for depression. The average low dose of intervention was 333.42 minutes and the average high dose (10studies) was 1005.00 minutes. Youth receiving low dose universal interventions had a significant mean SMG effect size of 0.27 (95% CI [0.13, 0.41], $p < .01$), and those receiving high dose universal interventions had a

significant SMG of 0.13 (95% CI [0.02, 0.24], $p < .05$). Direct comparison between low dose and high dose interventions, however, revealed no significant differences in mean effect size ($Z = 1.55$, $p > .05$).

Age. To test whether interventions were differentially effective according to participants' developmental level, mean effect sizes were compared based on the age of the samples that yielded them. Studies containing the necessary data were ranked according to average age of the participant sample, and the median average age was used to divide studies into "younger" and "older" participants. The mean ages of the study samples were then averaged across studies. For equality of comparison, the mean effect sizes produced by differing levels of intervention dose were once again limited to those obtained from evaluations of universal prevention studies.

There were eight studies of universal interventions for anxiety for which the average age of the sample was reported or could be obtained from study authors, and that had baseline and postintervention data. The average age of "younger" participants was 10.16 whereas with "older" participants the average age was 14.43. Younger youth receiving universal interventions had a significant mean SMG effect size of 0.32 (95% CI [0.14, 0.51], $p < .001$), and older youth receiving universal interventions had a significant SMG of 0.22 (95% CI [0.04, 0.40], $p < .05$). Direct comparison between younger and older youth receiving universal interventions, however, revealed no significant differences in mean effect size ($Z = 0.79$, $p > .05$).

There were eighteen studies of universal interventions for depression for which the average age of the sample was reported or could be obtained from study authors, and that had baseline and postintervention data. The average age of participants in the "younger" samples was 12.33 whereas with older participants the average age was 14.94. Younger subjects receiving universal interventions had a significant mean SMG effect size of 0.17 (95% CI [0.06, 0.29], $p <$

.01), and older youth receiving universal interventions had a significant SMG of 0.23 (95% CI [0.08, 0.37], $p < .01$). Direct comparison between younger and older youth receiving universal interventions, however, revealed no significant differences in mean effect size ($Z = 0.55$, $p > .05$).

Gender. To test whether participant gender moderated the effect of school-based CBT on reduction in disorder-related symptomatology, *Q*-tests were used to evaluate whether sex accounted for systematic variance in the effects of interventions. Within meta-analysis, *Q* tests – which are interpreted along the chi-square distribution – are statistical tests that perform a similar function as analyses of variance (ANOVAs) in that they allow for comparison of within- and between group variance to determine whether variability between groups exceeds the amount that would be expected to occur by chance alone (Borenstein et al., 2009). Only studies that presented data (e.g. means, standard deviations, and sample sizes) separately for males and females could be analyzed.

Two studies of school-based anxiety interventions reported data separately by gender of participants. For these studies, gender was not found to moderate changes in anxiety symptomatology from baseline to postintervention ($Q = .08$, $p > .05$).

Four studies of school-based depression interventions reported data separately by gender of participants. For these studies, gender was not found to moderate changes in depression symptomatology from baseline to postintervention ($Q = .30$, $p > .05$).

***Mediation Analyses*.** Are changes in the primary outcomes of anxiety and depression explained by changes in other constructs (e.g., self-esteem, hopelessness, explanatory style)? According to Baron and Kenny (1986) (see also Hopwood, 2007; Shrout & Bolger, 2002) a construct may serve as a mediating variable when it meets the following four criteria:

1) variations in levels of the independent variable significantly account for variation in the dependent variable.

2) variations in levels of the independent variable account for a significant amount of the variation in the hypothesized mediating variable.

3) variations in the hypothesized mediating variable significantly account for variations in the dependent variable.

4) After controlling for the relationships described in conditions 2 and 3, a previously significant relationship between the independent and dependent variables is no longer significant.

In the present meta-analytic review, however, neither self-esteem, hopelessness, nor explanatory/attributional style, were found to have improved significantly more from baseline to postintervention among youth receiving active interventions for depression as compared to controls. As such, the second necessary condition for mediation is not satisfied, thus precluding mediation analyses.

CHAPTER 4

DISCUSSION

This meta-analysis synthesized the available data from studies of school-based interventions for anxiety and depression and determined their effectiveness in reducing the symptomatology associated with these disorders. Analyses yielded support for the primary hypothesis: From 63 identified studies, mean effect size estimates for anxiety and depression interventions were statistically significant and were significantly greater than those obtained for the collection of comparison control groups. More specifically, summary effect sizes suggested that anxiety interventions were moderately effective, with a medium sized effect of 0.50, and that depression interventions were mildly effective, with a small-to-medium overall effect of 0.30. Fail-safe *N* calculations indicated that 108 anxiety intervention studies and 77 depression intervention studies, all with effect sizes of zero, would have to remain unidentified "in file drawers" to reduce these summary effect sizes to the criterion level of 0.10. As such, the present summary effects can be interpreted as relatively robust insofar as it is unlikely that such a number of additional unpublished studies all with an effect size of zero exist.

This study's hypotheses about effect size magnitude according to intervention classification level were supported. Figures 3-4 illustrate a stepwise pattern whereby treatment interventions yielded larger summary effect sizes than those produced by prevention programs, and prevention protocols that were based on some degree of participant risk (e.g. selective or indicated) yielded larger effect sizes than those delivered to all individuals regardless of risk (e.g. universal). This was true for both anxiety and depression interventions. These results are entirely consistent with what would be expected given the populations targeted by each level of intervention. For instance, very small effect sizes were yielded by universal interventions for

depression; given that the samples in such studies are often largely comprised of nonsymptomatic youth, it makes sense that mean symptom levels have limited room to decrease – reflecting the presence of "floor effects" in these studies. While understandable given the circumstances of their delivery, these small effect sizes produced by universal studies when combined with results about maintenance of gains (to be discussed) have raised important questions about the justifiability of such efforts (see below).

Contrary to one of this project's hypotheses, increasing duration of intervention was not associated with larger magnitude effect sizes. The various subsamples of anxiety and depression intervention studies, regardless of whether they were categorized as "low dose" or "high dose," all had significant mean effect sizes (albeit small), and there were no significant differences between the collections of "low dose" versus "high dose" studies. Such results could provide valuable information about the most cost-effective quantity of intervention that is needed for school-based interventions to effect short-term change. In this case, these findings can be viewed as promising, suggesting that protocols of a more compact nature may fare just as well as ones that are twice as long in terms of reducing symptomatology of anxiety and depression in youth.

Are intervention effects maintained over time? Findings from the present meta-analysis suggest that the answer to this question is "no." Based on the subsample of studies that explored outcomes approximately three months after postintervention, youth who received active interventions for anxiety continued to evidence significantly greater reductions in symptomatology from baseline as compared to controls, whereas no difference was found amongst depression intervention studies. The reverse was true by six month follow-up, at which point youth who had received interventions for depression demonstrated greater reductions in symptoms than controls, whereas no differences were found amongst anxiety intervention

studies. However, by 12-month follow-up, neither youth receiving anxiety interventions nor those in depression interventions exhibited significantly greater reductions in symptomatology from baseline than controls.

These results are consistent with prevailing attitudes in the field and expand upon existing knowledge. For instance, Spence and Shortt (2007) suggested that relatively brief universal prevention programs may be insufficient to yield lasting effects in the prevention of depression among youth. They concluded that endorsement of widespread school-based dissemination of such programs could be premature. Indeed, the results from the aggregated literature seem to indicate that although programs for internalizing disorders such as anxiety and depression may lead to short term reduction of symptoms by the end of an intervention, they do not confer the type of benefits that are necessary for yielding long-term superiority over no intervention at all.

Are the intervention effects different when a protocol is implemented by school staff as compared to members of the research team? In the present analysis, mean summary effect size findings suggested that there was no difference between these two groups of implementers in terms of the amount of reduction in symptomatology from baseline to postintervention – either for those studies exploring programs for anxiety or depression. Such results fall within the context of largely mixed findings throughout the literature regarding the general effectiveness of interventions when implemented by individuals other than research staff.

For instance, from among the studies included, one evaluation (Harnett & Dadds, 2004) devoted to evaluating a universal depression prevention program under real-world conditions with teachers as implementers found there to be no evidence for the effectiveness for those receiving the intervention as compared to a no-intervention control group. On the other hand, in

a trial (Barrett & Turner, 2001) that directly compared a teacher-led and a psychologist-led universal prevention intervention for anxiety, the authors found that both groups outperformed a monitoring control group, and reductions in anxious symptomatology were equitable. Additionally, Shochet and colleagues (2002) found there to be no significant differences between teachers and mental health professionals as group leaders for a universal prevention program either at posttreatment or follow-up for either boys or girls.

The present study reflects the most up-to-date analysis of the literature and reaches the conclusion that anxiety and depression interventions implemented by school staff yield outcomes equitable to those delivered by research staff. Such findings should be interpreted as good news, as long-term sustainability of school-based programs can only be achieved if interventions can be implemented effectively by individuals who have a consistent presence that spans multiple school years. However, it should also be noted that this particular analysis within the current project focused only on evaluations of universal prevention efforts, and thus may not be the case for other levels of intervention. For instance, Hunt and colleagues (2009) concluded that, although the school counselors and teachers implementing a prevention protocol achieved acceptable program fidelity, "for an indicated intervention…to be effective specialist mental health staff are needed to run it" (p.303). Thus, it may very well be that CBT interventions delivered to youth who evidence some degree of risk or are already manifesting significant symptoms require implementation by facilitators with more specialized training. While this may exclude individuals such as teachers, other school staff such as counselors, school social workers, or school psychologists could potentially deliver selective or indicated prevention, or perhaps even treatment – though further research is necessary to determine if such efforts could be feasible and effective.

An individual's tendency to interpret events as stable, negative, and global comprises the cognitive triad that has been demonstrated to underlie depression (Abramson, Alloy, & Metalsky, 1989). It makes sense, then, that all of the studies that included a measure of attributional style were school-based evaluations of interventions for depression. These investigations often sought to determine whether change in attributional style mediated the relationship between intervention status and reductions in depressive symptomatology. Results were mixed, however, in the conclusions drawn from study results. For instance, eight studies (e.g. Cardemil et al., 2002, Study 1; Cardemil et al., 2002; Chaplin et al., 2006; O'Kearney et al., 2006; O'Kearney et al., 2007; Pattison & Lynd-Stevenson, 2001; Quayle et al., 2001; Roberts et al., 2003) failed to find significantly more optimistic attributional style among intervention group participants than youth in control groups. One study (Horowitz et al., 2007) reported a nonsignificant trend for attributional style to partially mediate the relationship between intervention condition and depressive symptoms. Another study (Rooney et al., 2006) found that intervention group youth had a significantly more positive attributional style than those in the control group, but did not report on whether this relationship had a mediating effect on depression. Gillham and Seligman (1994) conservatively concluded that a significantly greater improvement in attributional style for intervention group youth did not mediate the significant reduction on depressive symptomatology, because the reduction in effect was not large enough to be considered significant according to recommendations set forth by Olkin and Finn (1990). However, they did report that a formerly significant relationship between intervention status and symptom change moved into the nonsignificant range after controlling for attributional style, which does satisfy the criteria for mediation established by Baron and Kenny (1986). Jaycox and colleagues (1994) at first failed to find any group differences between any composite scale scores

of explanatory style (e.g. positive, negative, overall composite). More detailed analysis of the specific dimensions of the negative composite scale, however, revealed that, by postintervention, participants in the prevention program were significantly less likely than control group youth to attribute negative events to stable, enduring causes, and that this change mediated the impact of the intervention on decrease of depressive symptomatology. In one report by Yu and Seligman (2002), it was found that change in explanatory style mediated the treatment effect on depressive symptoms. In the present analysis, the mean effect size for youth receiving interventions was not significantly different than for controls in terms of improving explanatory style, and thus no evidence for mediation relationships was obtained. This was also the case for self-esteem and hopelessness within evaluations of depression interventions.

An important issue for school systems seeking to offer mental health services is how to balance delivery of what is considered to be the most appropriate level of intervention with the potential risk for stigmatization. In terms of arguments for justification of intervention type, universal prevention efforts are endorsed as valuable because they provide skills and strategies for mental and emotional health that would be valuable to anyone. Advocates of selective prevention efforts, on the other hand, suggest that universal approaches are not cost effective – a major influencing factor that must be considered by school systems – and that recruitment on the basis of an established risk factor for disorder maximizes the usefulness of school-based programs. Still others who support indicated preventive efforts might claim that the most meaningful prevention is the one that can halt the progression of a clearly developing condition, the presence of which is evidenced by the existence of subclinical symptoms just below the threshold of disorder. Finally, there are those who do not discount the value of prevention, but

nonetheless are committed to implementing interventions that reflect “treatment” in the sense of providing services to only those students currently struggling with clinical level disorder.

One particular issue that often factors into this consideration is the potential for negative aspersions to be cast upon individuals receiving such services. For instance, in discussing the trade-offs between various intervention levels, Offord and colleagues (1998) list labeling and stigmatization as a disadvantage of targeted interventions while describing an absence of labeling and stigmatization as an advantage of universal interventions. Such a position is often endorsed by investigators. For instance, in describing universal prevention programs, Barrett, Lock, and Farrell (2005) suggest that “potential advantages involve reducing stigmatization” (p. 541). Other authors have echoed this position, claiming that universal preventive interventions “prevent the labeling effects of selective programs provided only to children ‘at risk’” (Quayle et al., 2001, p.195).

Indeed, the stigmatizing effect of categorizing youth as ‘at risk’ seems to be accepted as a foregone conclusion that many researchers ascribe to as commonsense, however as Rapee and colleagues (2006) have noted, there has been no empirical validation of this assumption. To address this, these authors conducted such an evaluation to determine whether students receiving an indicated intervention differed in their subjective experience from those receiving one that took a universal approach. Results found statistically significant differences in reported stigma, but the effect sizes were quite small and equated to very little absolute difference between youth across conditions. Perhaps more importantly, adolescents in the indicated preventive intervention reportedly significantly more satisfaction with the program they received than those in the universal intervention. Overall, Rapee and colleagues acknowledge that “there is little doubt that the indicated program was associated with some stigma” though they conclude that “this may be

a small price to pay for a mode of delivery that is potentially more satisfying for consumers and hence more sustainable" (p.175). These findings represent encouraging news suggesting that schools need not shy away from interventions based on risk status for fear of stigmatization; rather, they can and should implement the type of selective and indicated preventive efforts that were found to yield larger effect sizes than universal prevention in the current project.

The findings of the present meta-analysis should be viewed within the context of its limitations. First, the studies included were not limited to those classified as randomized controlled trials (RCTs), an approach that is sometimes taken when conducting reviews such as those published in the Cochrane Database (Higgins & Green, 2005). It is acknowledged that RCTs, due to their rigorous methodology, are likely to lead to the most reliable estimates of effects. However, given the challenges inherent to conducting school-based research (discussed above), a number of trials could not implement interventions according to RCT criteria. Thus, in an effort to examine the greatest number of distinct school-based evaluations, pilot studies or uncontrolled studies were included. Although it is believed that this decision helped to achieve the primary aim of maximizing the comprehensiveness of this review, it is recognized that such an approach may lead to the inclusion of individual study results that are less robust. Another limitation was the diminishing sample size of studies upon which secondary analyses were based. For instance, few studies collected data beyond one follow-up timepoint, which varied across studies. Similarly, few studies presented data separately by gender. As such, some findings are considerably limited by a small sample size of studies in manner akin to results from primary studies being limited by a small sample size of individual participants.

Synthesizing information from studies over the last 20 years, this meta-analysis updates and clarifies the data pertinent to implementation of mental health interventions within school

systems. One theme that emerged has to do with the absence of moderating variables: after aggregating data, effects on anxious and depressive symptomatology were not found to be moderated by age, gender, intervention dosage, or program implementer, nor were reductions in depression found to be mediated by associated changes in self-esteem, hopelessness, or attributional style. In terms of future directions, what is needed from research moving forward are studies that directly compare variations in key variables. One recommendation, echoing the exhortation of Calear and Christensen (2010) would be to "focus on the programs being utilised and what it is about their content and delivery style that make them more or less effective" (p.435).

Of central importance is the need to give serious consideration to the manner in which the "success" of protocols is measured, and whether we are falling short of assessing if such programs truly lead to meaningful change in the lives of youth. Many cognitive-behavioral interventions are aimed at modifying dysfunctional thought patterns in the hopes of leading to improved mood states. While this should continue to be a primary focus, it is recommended that future studies go beyond the standard paper-and-pencil questionnaires to determine whether, aside from a shift to more positive thinking, interventions are leading to true change in terms of *functional* outcome. For instance, are anxious youth actually engaging in less avoidant behaviors and are depressed youth actually participating in a higher rate of pleasurable activities? In essence, there remains a critical need to evaluate in an objective and data-based way whether improvements in internalizing symptomatology really translate into less of the interference that ultimately underlies anxious and depressive disorders.

The present meta-analysis found there to be no differences in results produced by interventions led by school staff as compared to those led by research staff, yet this is a

distinction that has often been referred to as being responsible for differential effects. Spence and Shortt (2007) observed that intervention protocol training and supervision during implementation vary widely across studies. Indeed, studies in the present review found implementer training to be inconsistently reported and vaguely described. Future research should examine whether school staff who receive extensive training implement an intervention with better effects than do those who simply read and follow a manual. Variations in the amount of ongoing supervision are also worthy of examination. A greater knowledge of these variables may help to explain the differences that have formerly been ascribed to various types of program implementers.

It is also recommended that future research more systematically report data on factors such as rates of attendance and program adherence. For variables such as these to be examined in future meta-analyses, it will be necessary for these data to be reported as part of standard practice in school-based research.

Finally, schools should consider further integration of mental health education and coping strategies into the curriculum. The present findings highlight that intervention effects are seldom maintained over the long term (see also Spence & Shortt, 2007). A hope of universal prevention efforts is that youth will learn skills that can be utilized down the road in a time of need. However, it may be that short-term (e.g. approximately 10 week) intervention is not sufficient to produce this result, as it seems clear that many programs are not producing the hoped-for lasting effects. Perhaps the time is right for a shift to a more developmental approach to prevention and intervention for youth anxiety and depression, and one that includes integrated interventions woven into the fabric of the regular curriculum. Moreover, if these efforts can be led by school staff, this would allow for the type of sustainability that would reflect a meaningful step in the

direction toward schools' achievement of meeting the growing mental health care needs of the children and adolescents.

The world that has become increasingly challenging for youth, but thankfully it has also become more accepting of the role of mental health services. It is time for schools to take on a more active role in providing developmental education in not only in reading, writing, and arithmetic, but also in how to cope with the difficulties that exist in life--the stress, anxiety, and depression that we now understand afflicts a significant proportion of today's youth.

REFERENCES CITED

References marked with an asterisk (*) indicate studies included in the meta-analysis.

Abramson, L. Y., Alloy, L. B., & Metalsky, G. I. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96,* 358-372.

Albano, A. M., Chorpita, B. F., & Barlow, D. H. (2003). Childhood anxiety disorders. In E. J. Mash, & Barkley, R. A. (Ed.), *Child psychopathology* (2nd ed., pp. 279-329). New York: Guilford Press.

Albee, G. W. & Gullotta, T. P. (1986). Facts and fallacies about primary prevention. *Journal of Primary Prevention, 6,* 207-218.

Andrews, G., Szabo, M., & Burns, J. (2002). Preventing major depression in young people. *British Journal of Psychiatry, 181*, 460-462.

Anglin, T. M., Naylor, K. E., & Kaplan, D. W. (1996). Comprehensive school-based health care: High school students' use of medical, mental health, and substance abuse services. *Pediatrics, 97*, 318-330.

APA Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48*, 3-24.

*Aune, T., & Stiles, T. C. (2009). Universal-based prevention of syndromal and sub-syndromal social anxiety: A randomized controlled study. *Journal of Consulting and Clinical Psychology, 77,* 867-8709.

Baron, R. M. & Kenny, D. A. (1986). The moderator–mediator variable distinction in

social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology,51,* 1173-1182.

Barrett, P. M., Dadds, Mark R., & Rapee, Ronald M. (1996). Family treatment of childhood anxiety: A controlled trial. *Journal of Consulting and Clinical Psychology, 64,* 333-342.

* Barrett, P. M., Farrell, L. J., Ollendick, T. H., & Dadds, M. (2006). Long-term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: An evaluation of the Friends program. *Journal of Clinical Child and Adolescent Psychology, 35,* 403-411.

* Barrett, P. M., Lock, S., & Farrell, L. J. (2005). Developmental differences in universal preventive intervention for child anxiety. *Clinical Child Psychology and Psychiatry, 10,* 539-555.

* Barrett, P. M., Moore, A. F., & Sonderegger, R. (2000). The FRIENDS Program for young former-Yugoslavian refugees in Australia: A pilot study. *Behaviour Change, 17,* 124-133.

* Barrett, P. M., Sonderegger, R., & Sonderegger, N.L. (2001). Evaluation of an anxiety-prevention and positive-coping program (FRIENDS) for children and adolescents of non-English-speaking background. *Behaviour Change, 18,* 78-91.

* Barrett, P. M., Sonderegger, R., & Xenos, S. (2003). Using FRIENDS to combat anxiety and adjustment problems among young migrants to Australia: A national trial. *Clinical Child Psychology and Psychiatry, 8,* 241-260.

* Barrett, P. & Turner, C. (2001). Prevention of anxiety symptoms in primary school children: Preliminary results from a universal school-based trial. *British Journal*

of Clinical Psychology, 40, 399-410.

Beardslee, W. R., Salt, P., Porterfield, K., Rothberg, P. C., van de Velde, Swatling, S., Hoke, L., Moilanen, D. L., & Wheelock, I. (1993). Comparison of preventive interventions for families with parental affective disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 254-263.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.

Bedell, J. R., & Roitzch, J. (1976). The effects of stress on state and trait anxiety in emotionally disturbed, normal, and delinquent children. *Journal of Abnormal Child Psychology, 4*, 173-177

Beidel, D. C., Turner, S. M., & Fink, C. M. (1996). Assessment of childhood social phobia: Construct, convergent, and discriminative validity of the Social Phobia and Anxiety Inventory for Children (SPAI-C). *Psychological Assessment, 8*, 235-240.

Beidel, D. C., Turner, S. M., Hamlin, K., & Morris, T. L. (2000). The Social Phobia and Anxiety Inventory for Children (SPAI-C): External and discriminative validity. *Behavior Therapy, 31*, 75-87.

Beidel, D. C., Turner, S. M., & Morris, T. L. (1995). A new inventory to assess childhood social anxiety and phobia: The Social Phobia and Anxiety Inventory for Children. *Psychological Assessment, 7*, 73-79.

* Bernstein, G. A., Bernat, D. H., Victor, A. M., & Layne, A. E. (2008). School-based interventions for anxious children: 3-, 6-, and 12-month follow-ups. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 1039-1047.

Bernstein, G. A., & Borchardt, C. M. (1991). Anxiety disorders of childhood and adolescence: A critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*, 519-532.

* Bernstein, G. A., Layne, A. E., Egan, E. A., & Tennison, D. M. (2005). School-based interventions for anxious children. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 1118-1127.

Birmaher, B., Brent, D. A., Chiapetta, L., Bridge, J., Monga, Suneeta, & Baugher, M. (1999). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): A replication study. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 1230-1236.

Birmaher, B., Khetarpal, S., Brent, D. Cully, M. et al. (1997). The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry, 36,* 545-553.

* Bokhorst, K., Goossens, F. A., & de Ruyter, P. A. (1995). Social anxiety at elementary school: The effects of a curriculum. *Educational Research, 37,* 87-94.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. West Sussex, UK: John Wiley & Sons

Brady, E. U., & Kendall, P. C. (1992). Comorbidity of anxiety and depression in children and adolescents. *Psychological Bulletin, 111,* 244-255.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

Brunwasser, S. M., Gillham, J. E., & Kim, E. S. (2009). A meta-analytic review of the

Penn Resiliency Program's effect on depressive symptoms. *Journal of Consulting and Clinical Psychology, 77,* 1042-1054.

Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M. Z., & Erkanli, A. (1995). DataWatch: Children's mental health service use across service sectors. *Health Affairs, 14,* 147-159.

Calear, A. L., & Christensen, H. (2010). Systematic review of school-based prevention and early intervention programs for depression. *Journal of Adolescence, 33,* 429-438.

* Calear, A. L., Christensen, H., Mackinnon, A., Griffiths, K. M., & O'Kearney, R. (2009). The YouthMood project: A cluster randomized controlled trial of an online cognitive behavioral program with adolescents. *Journal of Consulting and Clinical Psychology, 77,* 1021-1032.

* Cardemil, E. V., Reivich, K. J., Beevers, C. G., Seligman, M. E. P., & James, J. (2007). The prevention of depressive symptoms in low-income, minority children: Two-year follow-up. *Behaviour Research and Therapy, 45,* 313-327.

* Cardemil, E. V., Reivich, K. J., & Seligman, M. E. P. (2002). The prevention of depressive symptoms in low-income minority middle school students. *Prevention and Treatment, 5,* 110-126.

* Chaplin, T. M., Gillham, J. E., Reivich, K., Elkon, A. G. L., Freres, D. R., Winder, B., et al. (2006). Depression prevention for early adolescent girls: A pilot study of all girls versus co-ed groups. *The Journal of Early Adolescence, 26,* 110-126.

Chaplin, W. F. (1984). State Trait Anxiety Inventory. In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques* (Vol. 1, pp. 626-632). Kansas City, MO: Test Corporation of America.

Chavira, D. A., Stein, M. B., Bailey, K., & Stein, M. T. (2004). Child anxiety in primary care: Prevalent but untreated. *Depression and Anxiety, 20,* 155-164.

Chambless, D. L. & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66,* 7-18.

Chorpita, B. F. (2003). The frontier of evidence-based practice. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 42-59). New York: Guildford.

Clarke, G. N., Hawkins, W., Murphy, M., Sheeber, L. B., Lewinsohn, P. M., & Seeley, J. R. (1995). Targeted prevention of unipolar depressive disorder in an at-risk sample of high school adolescents: A randomized trial of a group cognitive intervention. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 312-321.

* Clarke, G. N., Hawkins, W., Murphy, M., Sheeber, L. (1993). School-based primary prevention of depressive symptomatology in adolescents: Findings from two studies. *Journal of Adolescent Research, 8,* 183-204.

Clarke, G. N., Hornbrook, M., Lynch, F., Polen, M., Gale, J., Beardslee, W., O'Connor, E., & Seeley, J. (2001). A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. *Archives of General Psychiatry, 58,* 1127-1134.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cole, D. A., Peeke, L. G., Martin, J. M., Truglio, R., & Seroczynsky, A. D. (1998). A longitudinal look at the relation between depression and anxiety in children and adolescents. *Journal of Consulting and Clinical Psychology, 66,* 451-460.

Collins, K. A. & Dozois, D. J. A. (2008). What are the active ingredients in preventative interventions for depression? *Clinical Psychology: Science and Practice, 15*, 313-330.

Comer, J. S., Furr, J. M., Beidas, R. S., Weiner, C. L., & Kendall, P. C. (2008). Children and terrorism-related news: Training parents in coping and media literacy. *Journal of Consulting and Clinical Psychology, 76*, 568-578.

Comer, J. S. & Kendall, P. C. (2007). Terrorism: The psychological impact on youth. *Clinical Psychology: Science and Practice, 14*, 178-212.

Commission on Chronic Illness. (1957). *Chronic Illness in the United States* (Vol. 1). Cambridge, MA: Harvard University Press.

Commonwealth Department of Health and Aged Care. (2000). *Promotion, prevention and early intervention for mental health – A monograph.* Canberra: Mental Health and Special Programs Branch, Commonwealth Department of Health and Aged Care.

Compton, S. N., March, J. S., Brent, D., Albano, A. M., Weersing, V. R., & Curry, J. (2004). Cognitive-behavioral psychotherapy for anxiety and depressive disorders in children and adolescents: An evidence-based medicine review. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 930-959.

Cook, D. J., Guyatt, G. H., & Ryan, G. (1993). Should unpublished data be included in meta-analyses? Current convictions and controversies. *Journal of the American Medical Association, 269*, 2749-2753.

* Cooley, M. R., Boyd, R. C., & Grados, J. J. (2004). Feasibility of an anxiety preventive intervention for community violence exposed African-American children. *The Journal of Primary Prevention, 25*, 105-123.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York: Russel Sage Foundation.

Cooper, H.M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin, 87,* 442-449.

Costello, J. E., Egger, H. L., & Angold, A. (2004). Developmental epidemiology of anxiety disorders. In T. H. Ollendick, & March, J. S. (Eds.), *Phobic and anxiety disorders in children and adolescents: A clinician's guide to effective psychosocial and pharmacological interventions* (pp. 61-91). New York: Oxford University Press.

Costello, J. E., Egger, H. L., & Angold, A. (2005). 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 972-986.

Costello, J.E., Mustillo, S., Erkanli, A., Keller, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry, 60*, 837–844.

Council of Australian Governments (COAG). National Action Plan for Mental Health 2006-2011. (accessed February 2009) http://www.coag.gov.au/reports/docs/AHMC_COAG_mental_health.pdf

Cukrowicz, K. C., White, B. A., Reitzel, L. R., Burns, A. B., Driscoll, K. A., Kemper, T. S., et al. (2005). Improved treatment outcome associated with the shift to empirically supported treatments in a graduate training clinic. *Professional Psychology: Research and Practice, 36*, 330-337.

*Curtis, S. E. (1992). *Cognitive–behavioral treatment of adolescent depression.*

Unpublished doctoral dissertation, Utah State University, Logan.

* Dadds, M. R., Holland, D. E., Laurens, K. R., Mullins, M., Barrett, P. M., & Spence, S. H. (1999). Early intervention and prevention of anxiety disorders in children: Results at 2-year follow-up. *Journal of Consulting and Clinical Psychology, 67,* 145-150.

* Dadds, M. R., Spence, S. H., Holland, D. E., Barrett, P. M., & Laurens, K. R. (1997). Prevention and early Intervention for anxiety disorders: A controlled trial. *Journal of Consulting and Clinical Psychology, 65*, 627-635.

Dalley, M. B., Bolocofsky, D. N., Alcorn, M. B., & Baker, C. (1992). Depressive symptomatology, attributional style, dysfunctional attitude, and social competency in adolescents with and without learning disabilities. *School Psychology Review, 21*, 444-458.

Dekking, Y. M. (1983). *SAS-K Sociale Angst Schaal voor Kinderen* (*The Social Anxiety Scale for Children*). Lisse: Swets and Zeitlinger.

Durlak, J. A., & Lipsey, M. W. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology, 19,* 291-332.

Earls, R., Robins, L. N., Stiffman, A. R., & Powell, J. (1989). Comprehensive healthcare for high-risk adolescents: An evaluation study. *American Journal of Public Health, 79,* 999-1005.

*Ettelson, R. G. (2003). *The treatment of adolescent depression.* Unpublished doctoral dissertation, Illinois State University, Normal.

Evans, S. W. & Weist, M. D. (2004). Implementing empirically supported treatment in the schools: What are we asking? *Clinical Child and Family Psychology Review, 7,* 263-267.

Fauber, R., Forehand, R., Long, N., Burke, M., & Faust, J. (1987). The relationship of young adolescent Children's Depression Inventory (CDI) scores to their social and cognitive functioning. *Journal of Psychopathology and Behavioral Assessment, 9*, 161-172.

Finch, A. J., Saylor, C. F., & Edwards, G. L. (1985). Children's Depression Inventory: Sex and grade norms for normal children. *Journal of Consulting and Clinical Psychology, 53*, 424-425.

Finch, A. J., Saylor, C. F., Edwards, G. L., & McIntosh, J. A. (1987). Children's Depression Inventory: Reliability over repeated administrations. *Journal of Clinical Child Psychology*, *16*, 339-341.

Flaherty, L. T., Weist, M. D., & Warner, B. S. (1996). School-based mental health services in the United States: History, current models and needs. *Community Mental Health Journal, 32*, 341-352.

Flannery-Schroeder, E., Choudhury, M. S., & Kendall, P. C. (2005). Group and individual cognitive-behavioral treatments for youth with anxiety disorders: 1-year follow-up. *Cognitive Therapy and Research, 29,* 253-259.

Flannery-Schroeder, E. C. & Kendall, P. C. (2000). Group and individual cognitive-behavioral treatments for youth with anxiety disorders: A randomized clinical trial. *Cognitive Therapy and Research, 24*, 251-278.

* Gallegos, J. (2008). *Preventing childhood anxiety and depression: Testing the effectiveness of a school-based program in México*. Unpublished doctoral dissertation, The University of Texas at Austin, Austin.

Georgiades, K., Lewinsohn, P. M., Monroe, S. M., & Seeley, J. R. (2006). Major

depressive disorder in adolescence: The role of subthreshold symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*, 936-944.

* Gillham, J. E., & Seligman, M. E. P. (1994). *Preventing depressive symptoms in school children.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

* Gillham, J. E., & Reivich, K. J. (1999). Prevention of depressive symptoms in schoolchildren: A research update. *Psychological Science, 10*, 461-462.

* Gillham, J. E., Reivich, K. J., Freres, D. R., Chaplin, T. M., Shatté, A. J., Samuels, B., Elkon, A. G., Litzinger S., Lascher, M., Gallop, R., & Seligman, M. E. P. (2007). School-based prevention of depressive symptoms: A randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. *Journal of Consulting and Clinical Psychology, 75*, 9-19.

* Gillham, J. E., Reivich, K. J., Freres, D. R., Lascher, M., Litzinger, S., Shatté, A. & Seligman, M. E. P. (2006). School-based prevention of depression and anxiety symptoms in early adolescence: A pilot of a parent intervention component. *School Psychology Quarterly, 21*, 323-348.

* Gillham, J. E., Reivich, K. J., Jaycox, L. H., & Seligman, M. E. P. (1995). Prevention of depressive symptoms in schoolchildren: Two-year follow-up. *Psychological Science, 6*, 343-351.

Ginsburg, G. S., Becker, K. D., Kingery, J. N., & Nichols, T. (2008). Transporting CBT for childhood anxiety disorders into inner-city school-based mental health clinics. *Cognitive and Behavioral Practice, 15*, 148-158.

* Ginsburg, G. S., & Drake, K. L. (2002). School-based treatment for anxious African-

American adolescents: A controlled pilot study. *Journal of the American Academy of Child and Adolescent Psychiatry, 41,* 768-775.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5,* 3-8.

Gordon, R. (1983). An operational classification of disease prevention. *Public Health Reports; 98,* 107-109.

Gordon, R. (1987). An operational classification of disease prevention. In J.A. Steinberg & M.M. Silverman (Eds.), *Preventing mental disorders* (pp. 20-26).Rockville, MD: Department of Health and Human Services.

Gotlib, I. H., Lewinsohn, P. M., & Seeley, J. R. (1995). Symptoms versus a diagnosis of depression: Differences in psychosocial functioning. *Journal of Consulting and Clinical Psychology, 63,* 90-100.

* Hains, A. A. (1992). Comparison of cognitive-behavioral stress management techniques with adolescent boys. *Journal of Counseling & Development, 70,* 600-605.

* Hains, A. A. (1994). The effectiveness of a school-based, cognitive-behavioral stress management program with adolescents reporting high and low levels of emotional arousal. *School Counselor, 42,* 114-125.

* Hains, A. A., & Szyjakowski, M. (1990). A cognitive stress-reduction intervention program for adolescents. *Journal of Counseling Psychology, 37,* 79-84.

Hammen, C., & Rudolph, R. D. (2003). Childhood mood disorders. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (2nd ed., pp. 233-278). New York: Guilford Press.

Hannan, A. P., Rapee, R. M., & Hudson, J. L. (1995). The prevention of depression in children: A pilot study. *Behaviour Change, 17,* 78-83.

* Harnett, P. H. & Dadds, M. R. (2004). Training school personnel to implement a universal school-based prevention of depression program under real-world conditions. *Journal of School Psychology, 42,* 343-357.

Heckel, L., Clarke, A., Barry, R., McCarthy, R., & Selikowitz, M. (2009). The relationship between divorce and the psychological well-being of children with ADHD: Differences in age, gender, and subtype. *Emotional & Behavioural Difficulties, 14*, 49-68.

Hedges, L.V. (1982). Estimation of effect size from a series of independent studies. *Psychological Bulletin, 92*, 490-499.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486-504.

Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65,* 821-833.

Henggeler, S. W., Schoenwald, S. K., & Pickrel, S. G. (1995). Multisystemic Therapy: Bridging the gap between university- and community-based treatment. *Journal of Consulting and Clinical Psychology, 63*, 709-717.

Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [updated May 2005]. In: The Cochrane Library, Issue 3, 2005. Chichester, UK: John Wiley & Sons, Ltd.

Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557-560.

Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 683-687.

Hoffart, A., & Martinsen, E. W. (1991). Causal attributions in clinical subtypes of depression: A longitudinal study of inpatients. *Journal of Psychopathology and Behavioral Assessment, 13,* 241-256.

Hoffman, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78,* 169-183.

Hopwood, C. J. (2007). Moderation and mediation in structural equation modeling: Applications for early intervention research. *Journal of Early Intervention, 29,* 262-272.

Horowitz, J. L., & Garber, J. G. (2006). The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 74,* 401-415.

* Horowitz, J. L., Garber, J. G., Ciesla, A., Young, J. F., & Mufson, L. (2007). Prevention of depressive symptoms in adolescents: A randomized trial of cognitive-behavioral and interpersonal prevention programs. *Journal of Consulting and Clinical Psychology,75,* 693-706.

* Hunt, C., Andrews, G., Crino, R., Erskine, A., & Sakashita, C. (2009). Randomized

controlled trial of any early intervention programme for adolescent anxiety disorders. *Australian and New Zealand Journal of Psychiatry, 43,* 300-304.

Hunter, J.E., & Schmidt, F.L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Ialongo, N., Edelsohn, G., Werthamer-Larsson, L., Crockett, L., & Kellam, S. (1994). The significance of self-reported anxious symptoms in first-grade children. *Journal of Abnormal Child Psychology, 22*, 441-455.

Ialongo, N., Edelsohn, G., Werthamer-Larsson, L., Crockett, L., & Kellam, S. (1995). The significance of self-reported anxious symptoms in first grade children: Prediction to anxious symptoms and adaptive functioning in fifth grade. *Journal of Child Psychology and Psychiatry, 36*, 427-437.

Ishikawa, S., Okajima, I., Matsuoka, H., & Sakano, Y. (2007). Cognitive behavioural therapy for anxiety disorders in children and adolescents: A meta-analysis. *Child and Adolescent Mental Health, 12*, 164-172.

James, A. A. C. J., Soler, A., Weatherall, R. R. W. (2005). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews,* Issue 4. Art. No.: CD004690. DOI: 10.1002/14651858.CD004690.pub2.

* Jaycox, L. H., Reivich, K. J., Gillham, J., & Seligman, M. E. P. (1994). Prevention of depressive symptoms in school children. *Behaviour Research and Therapy, 32,* 801-816.

Jaycox, L. H., Stein, B. D., Paddock, S., Miles, J. N. V., Chandra, A., & Meredith, L. S., et al. (2009). Impact of teen depression on academic, social, and physical functioning. *Pediatrics, 124*, e596-e605.

* Kahn, J. S., Kehle, T. J., Jenson, W. R., & Clark, E. (1990). Comparison of cognitive-behavioral, relaxation, and self-modeling interventions for depression among middle-school students. *School Psychology Review, 19,* 196-211.

Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 62,* 100-110.

Kendall, P. C. (2006). Guiding theory for therapy with children and adolescents. In P.C. Kendall (Ed.). *Child and adolescent therapy: Cognitive-behavioral procedures* (3rd ed.; pp. 3-30). New York, NY: Guilford Press.

Kendall, P. C. & Beidas, R.S. (2007). Smoothing the trail for dissemination of evidence-based practices for youth: Flexibility within fidelity. *Professional Psychology: Research and Practice, 38,* 13-20.

Kendall, P. C., Flannery-Schroeder, E., Panichelli-Mindel, S. M., Southam-Gerow, M., Henin, A., & Warman, M. (1997). Therapy for youths with anxiety disorders: A second randomized clinical trial. *Journal of Consulting and Clinical Psychology, 65,* 366-380.

Kendall, P.C. & Hedtke, K.A. (2006a). *Cognitive-behavioral therapy for anxious children: Therapist manual* (3rd ed.). Ardmore, PA: Workbook Publishing. www.WorkbookPublishing.com

Kendall, P.C. & Hedtke, K.A. (2006b). *The Coping Cat Workbook* (2nd ed.). Ardmore, PA: Workbook Publishing. www.WorkbookPublishing.com

Kendall, P. C., Hudson, J. L., Gosch, E., Flannery-Schroeder, E., & Suveg, C. (2008). Cognitive-behavioral therapy for anxiety disordered youth: A randomized clinical trial evaluating child and family modalities. *Journal of Consulting and Clinical Psychology, 76,* 282-297.

Kendall, P. C. & Norton-Ford, J. D. (1982). *Clinical psychology: Scientific and professional dimensions*. New York, NY: Wiley.

Kendall, P. C., Safford, S., Flannery-Schroeder, E., & Webb, A. (2004). Child anxiety treatment: Outcomes in adolescence and impact on substance abuse and depression at 7.4 year follow-up. *Journal of Consulting and Clinical Psychology, 72,* 276-287.

Kendall, P. C., & Southam-Gerow, M. A. (1995) Issues in the transportability of treatment: The case of anxiety disorders in youths. *Journal of Consulting and Clinical Psychology, 63,* 702-708.

* Kiselica, M. S., Baker, S. B., Thomas, R. N., & Reedy, S. (1994). Effects of stress inoculation training on anxiety, stress, and academic performance among adolescents. *Journal of Counseling Psychology, 41,* 335-342.

Kovacs, M. (1985). The Children's Depression Inventory (CDI). *Psychopharmacology Bulletin, 21,* 995–1124.

Kovacs, M. (1992). *Children's Depression Inventory Manual*. New York: Multi-Health Systems Inc.

Kovacs, M. (1996). The course of childhood-onset depressive disorders. *Psychiatric Annals, 26,* 326-330.

* Kowalenko, N., Rapee, R.M., Simmons, J., Wignall, A., Hoge, R., Whitefield, K., Starling, J., Stonehouse, R., & Baillie, A. J. (2005). Short-term effectiveness of a school-based early intervention program adolescent depression. *Clinical Child Psychology and Psychiatry, 10,* 493-507.

Kutash, K., Duchnowski, A. J. & Lynn, N, (2006). *School-based mental health: An empirical guide for decision-makers.* Tampa, FL: University of South Florida, The Louis de la Parte Florida Mental Health Institute, Department of Child & Family Studies., Research and Training Center for Children's Mental Health.

La Greca, A. M. (1998). *Manual for the Social Anxiety Scales for Children and Adolescents.* Available from the author: Department of Psychology, PO Box 249229, University of Miami, Coral Gables, FL 33124.

La Greca, A. M., & Stone, W. L. (1993). The Social Anxiety Scale for Children-Revised: factor structure and concurrent validity. *Journal of Clinical Child Psychology, 22*, 17–27.

Lewinsohn, P. M. & Clarke, G. N. (1999). Psychosocial treatments for adolescent depression. *Clinical Psychology Review, 19*, 329-342.

Lewinsohn, P. M., Clarke, G. N., Hops, H., & Andrews, J. (1990). Cognitive-behavioral group treatment of adolescent depression. *Behavior Therapy, 21*, 385-401.

* Liddle, B., & Spence, S. H. (1990). Cognitive-behaviour therapy with depressed primary school children: A cautionary note. *Behavioural Psychotherapy, 18*, 85-102.

Lipsey, M. W., & Wilson, D. B. (2000). Practical meta-analysis. *Applied Social Research Methods Series* (Vol. 49). Thousand Oaks, CA: Sage Publications.

* Lock, S. & Barrett, P. M. (2003). A longitudinal study of developmental differences in universal preventive intervention for child anxiety. *Behaviour Change, 20*, 183-199.

Logan, D. E. & King, C. A. (2002). Parental identification of depression and mental health service use among depressed adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*, 296-304.

* Lowry-Webster, H. M., Barrett, P. M., & Dadds, M. R. (2001). A universal prevention trial of anxiety and depressive symptomatology in childhood: Preliminary data from an Australian Study. *Behaviour Change, 18*, 36-50.

* Lowry-Webster, H. M., Barrett, P. M., & Lock, S. (2003). A universal prevention trial

of anxiety symptomatology during childhood: Results at 1-year follow-up. *Behaviour Change, 20,* 25-43.

March, J. S. (1997). *The Multidimensional Anxiety Scale for Children*. North Tonawanda, NY: Multi-Health Systems.

March, J. S., Parker, J. D. A., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 554-565.

* Masia, C. L., Klein, R. G., Storch, E. A., Corda, B. (2001). School-based behavioral treatment for social anxiety disorder in adolescents: Results of a pilot study. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 780-786.

* Masia-Warner, C., Klein, R. G., Dent, H. C., Fisher, P. H., Alvir, J., Albano, A. M., & Guardino, M. (2005). School-based intervention for adolescents with social anxiety disorder: Results of a controlled study. *Journal of Abnormal Child Psychology, 33,* 707-722.

McCabe, O. L. (2004). Cross the quality chasm in behavioral health care: The role of evidence-based practice. *Professional Psychology: Research and Practice, 35,* 571-579.

McLean, P. D. & Hakstian, R. A. (1979). Clinical depression: Comparative efficacy of outpatient treatments. *Journal of Consulting and Clinical Psychology, 47,* 818-836.

McLoone, J., Hudson, J. L., & Rapee, R. M. (2006). Treating anxiety disorders in a school setting. *Education and Treatment of Children, 29,* 219-242.

Merry, S., McDowell, H., Hetrick, S., Bir J., & Muller, N. (2004). Psychological and/or

educational interventions for the prevention of depression in children and adolescents. *The Cochrane Database of Systematic Reviews* 2004, Issue 2. Art. No.: CD003380.pub2. DOI:10.1002/14651858.CD003380.pub2.

* Merry, S., McDowell, H., Wild, C. J., Bir, J., Cunliffe, R. (2004). A randomized placebo-controlled trial of a school-based depression prevention program. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 538-547.

Messner, S. F., Bjarnason, T., Raffalovich, L. E., & Robinson, B. K. (2006). Nonmarital fertility and the effects of divorce rates on youth suicide rates. *Journal of Marriage and Family, 68*, 1105-1111.

* Mifsud, C. & Rapee, R. M. (2005). Early intervention for childhood anxiety in a school setting: Outcomes for an economically disadvantaged population. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 996-1004.

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., Stroup, D. F., for the QUOROM Group (1999). Improving the quality of reports of meta-analysis of randomized controlled trials: the QUOROM statement. *Lancet, 354*, 1896-1900.

Morris, T. L. & Masia, C. L. (1998). Psychometric evaluation of the Social Phobia and Anxiety Inventory for Children: Concurrent validity and normative data. *Journal of Clinical Child Psychology, 27*, 452-458.

Moses, L. E., Mostellar, F., & Buehler, J. H. (2002). Comparing results of large clinical trials to those of meta-analyses. *Statistics in Medicine, 21*, 793-800.

* Mostert, J., & Loxton, H. (2008). Exploring the effectiveness of the FRIENDS program in reducing anxiety symptoms among South African children. *Behaviour Change, 25*, 85-96.

* Muris, P., Bogie, N., & Hoogsteder, A. (2001). Effects of an early intervention group program for anxious and depressed adolescents: A pilot study. *Psychological Reports, 88,* 481-482.

* Muris, P., Mayer, B., Bartelds, E., Tierney, S., & Bogie, N. (2001). The revised version of the Screen for Child Anxiety Related Emotional Disorders (SCARED-R): Treatment sensitivity in an early intervention trial for childhood anxiety disorders. *British Journal of Clinical Psychology, 40,* 323-336.

Mrazek, P. J. & Haggerty, R. J. (Eds.) *Reducing risks for mental disorders: Frontiers for prevention research.* Washington, DC: National Academy Press.

Mufson, L. H., Dorta, K. P., Olfson, M., Weissman, M. M., & Hoagwood, K. (2004). Effectiveness research: Transporting interpersonal psychotherapy for depressed adolescents (IPT-A) from the lab to school-based health clinics. *Clinical Child and Family Psychology Review, 7,* 251-261.

Myers, K. & Winters, N. C. (2002). Ten-year review of rating scales. II: Scales for internalizing disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 41,* 634-659.

National Research Council. (1992). *Combining information: Statistical issues and opportunities for research*. Washington, DC: National Academy of Sciences Press.

Neil, A. L., & Christensen, H. (2007). Australian school-based prevention and early intervention programs for anxiety and depression: A systematic review. *Medical Journal of Australia, 186,* 305-208.

Neil, A. L., & Christensen, H. (2009). Efficacy and effectiveness of school-based

prevention and early intervention programs for anxiety. *Clinical Psychology Review, 29*, 208-215.

New Freedom Commission on Mental Health. (2003) *Achieving the promise: Transforming mental health care in America.* Final Report. DHHS Pub. No. SMA-03-3832. Rockville, MD.

Offord, D. R., Kraemer, H. C., Kazdin, A. E., Jensen, P. S., & Harrington, R. (1998). Lowering the burden of suffering from child psychiatric disorder: Trade-offs among clinical, targeted, and universal interventions. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*, 686-694.

* O'Kearney, R., Gibson, M., Christensen, H., & Griffiths, K. M. (2006). Effects of a cognitive-behavioural internet program on depression, vulnerability to depression and stigma in adolescent males: A school-based controlled trial. *Cognitive Behaviour Therapy, 35*, 43-54.

* O'Kearney, R., Kang, K., Christensen, H., & Griffiths, K. (2009). A controlled trial of a school-based internet program for reducing depressive symptoms in adolescent girls. *Depression and Anxiety, 26*, 65-72.

Olkin, I. & Finn, J. (1990). Testing correlated correlations. *Psychological Bulletin, 108*, 330-333.

Ollendick, T. H., King, N. J., & Chorpita, B. F. (2006). Empirically supported treatments for children and adolescents. In P.C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures (3rd ed.),* (pp. 492-520). New York: Guildford Press.

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157-159.

Owens, J. S., & Murphy, C. E. (2004). Effectiveness research in the context of school-based mental health. *Clinical Child and Family Psychology Review, 7*, 195-209.

* Pattison, C., & Lynd-Stevenson, R. M. (2001). The prevention of depressive symptoms in children: The immediate and long-term outcomes of a school-based program. *Behaviour Change, 18*, 92-102.

Pincus, D. B., & Friedman, A. G. (2004). Improving children's coping with everyday stress: Transporting treatment interventions to the school setting. *Clinical Child and Family Psychology Review, 7*, 223-240.

* Pössel, P., Horn, A.B., Groen, G., & Hautzinger, M. (2004). School-based prevention of depressive symptoms in adolescents: A 6-month follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 43, 1003-1010.*

* Pössel, P., Seemann, S., & Hautzinger, M. (2008). Impact of comorbidity in prevention of adolescent depressive symptoms. *Journal of Counseling Psychology, 55*, 106-117.

* Quayle, D., Dziurawiec, S., Roberts, C., Kane, R., & Ebsworthy, G. (2001). The effect of an optimism and lifeskills program on depressive symptoms in preadolescence. *Behaviour Change, 18*, 194-203.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385-401.

Ralph, A., & Nicholson, L. (1995). Teaching coping skills to depressed adolescents in high school settings. *Behaviour Change, 12*, 175-190.

Rapee, R.M., Wignall, A., Sheffield, J., Kowalenko, N., Davis, A., McLoone, J. et al. (2006). Adolescents' reactions to universal and indicated prevention programs for depression: Perceived stigma and consumer satisfaction. *Prevention Science, 7*, 167-177.

Reivich, K. (1996). *The prevention of depressive symptoms in adolescents*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Reynolds, C. R. (1982). Convergent and divergent validity of the Revised Children's Manifest Anxiety Scale. *Educational and Psychological Measurement, 42*, 1205-1212.

Reynolds, C. R. & Paget, K. D. (1983). National normative and reliability data for the Revised Children's Manifest Anxiety Scale. *School Psychology Review, 12*, 324-336.

Reynolds, C. R. & Richmond, O. B. (1978). What I think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology, 6*, 271-280.

Reynolds, W. M. (1986). *Reynolds Adolescent Depression Scale*. Odessa, FL: Psychological Assessment Resources Inc.

Reynolds, W. M. (1989). *Reynolds Child Depression Scale*. Odessa, FL: Psychological Assessment Resources.

Reynolds, W. M. (1990). Depression in children and adolescents: Nature, diagnosis, assessment, and treatment. *School Psychology Review, 19*, 158-173.

Reynolds, W. M. (1992). Depression in children and adolescents. In W. M. Reynolds (Ed.) *Internalizing disorders in children and adolescents*. New York: Wiley.

Roberts, R. E., Attkisson, C. C., & Rosenblatt, A. (1998). Prevalence of psychopathology among children and adolescents. *American Journal of Psychiatry, 155*, 715-725.

* Roberts, C., Kane, R., Bishop, B., Matthews, H., & Thomson, H. (2004). The prevention of depressive symptoms in rural school children: A follow-up study. *International Journal of Mental Health Promotion, 6*, 4-16.

* Roberts, C., Kane, R., Thomson, H., Bishop, B., & Hart, B. (2003). The prevention of depressive symptoms in rural school children: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 71*, 622-628.

Rones, M., & Hoagwood, K. (2003). School-based mental health services: A research review. *Clinical Child and Family Psychology Review, 3*, 223-241.

* Rooney, R., Roberts, C., Kane, R., Pike, L., Winsor, A., White, J., et al. (2006). The prevention of depression in 8- to 9-year-old children: A pilot study. *Australian Journal of Guidance and Counseling, 16*, 76-90.

* Rose, H., Miller, L., & Martinez, Y. (2009). "FRIENDS for Life": The results of a resilience-building, anxiety-prevention program in a Canadian elementary school. *Professional School Counseling, 12*, 400-407.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638-641.

Rosenthal, R. (1984). Meta-analytic procedures for social research. *Applied Social Research Methods Series* (Vol. 6). Beverly Hills, CA: Sage Publications.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. London: Sage Publications.

Rosenthal, R. (1993). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183-192.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59-82.

Schoenwald, S. K., Chapman, J. E., Kelleher, K., Hoagwood, K. E., Landsverk, J., Stevens, J, et

al. (2008). A survey of the infrastructure for children's mental health services: Implications for the implementation of empirically supported treatments (ESTs). *Administration and Policy in Mental Health and Mental Health Services Research, 35,* 84-97.

Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services, 52,* 1190-1197.

Shadish, W. R., & Sweeney, R. B. (1991). Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. *Journal of Consulting and Clinical Psychology, 59,* 883-893.

Shaw, B. F. (1977). Comparison of cognitive therapy and behavior therapy in the treatment of depression. *Journal of Consulting and Clinical Psychology, 45,* 543-551.

* Sheffield, J. K., Spence, S. H., Rapee, R. M., Kowalenko, N., Wignall, A., Davis, A., & McLoone, J. (2006). Evaluation of universal, indicated, and combined cognitive-behavioral approaches to the prevention of depression among adolescents. *Journal of Consulting and Clinical Psychology, 74,* 66-79.

* Shochet, I. M., Dadds, M. R., Holland, D., Whitefield, K., Harnett, P. H., & Osgarby, S. M. (2001). The efficacy of a universal school-based program to prevent adolescent depression. *Journal of Clinical Child Psychology, 30,* 303-315.

* Shochet, I. M., Montague, R., & Dadds, M. (2002). *Preventing depressive symptoms in adolescence with sustainable resources: Evaluation of a school-based effectiveness trial*. Unpublished manuscript.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and studies: New procedures and recommendations. *Psychological Methods, 7,* 422-445.

Silverman, W. K., Kurtines, W. M., Gins/burg, G. S., Weems, C. F., Lumpkin, P. W., & Carmichael, D. H. (1999). Treating anxiety disorders in children with group cognitive-behavioral therapy: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 67,* 995-1003.

Silverman, W. K., Pina, A. A., & Viswesvaran, C. (2008). Evidence-based psychosocial treatments for phobic and anxiety disorders in children and adolescents. *Journal of Clinical Child & Adolescent Psychology, 37,* 105-130.

Spence, S. H. (1997). Structure of anxiety symptoms in children: A confirmatory factor-analytic study. *Journal of Abnormal Psychology, 106,* 280–297.

Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy, 36,* 545-566.

* Spence, S. H., Sheffield, J. K., Donovan, C. L. (2003). Preventing adolescent depression: An evaluation of the problem solving for life program. *Journal of Consulting and Clinical Psychology, 71,* 3-13.

* Spence, S. H., Sheffield, J. K., & Donovan, C. L. (2005). Long-term outcome of a school-based, universal approach to prevention of depression in adolescents. *Journal of Consulting and Clinical Psychology, 73,* 160-167.

Spence, S. H., & Shortt, A. L. (2007). Research review: Can we justify the widespread dissemination of universal, school-based interventions for the prevention of depression among children and adolescents? *Journal of Child Psychology and Psychiatry, 48,* 526-542.

Spielberger, C. D. (1973). *Manual for the State–Trait Anxiety Inventory for Children.* Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Gorsuch, R. L., & Lushene R. E. (1970). *State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

* Stallard, P., Simpson, N., Anderson, S., Carter, T., Osborn, C., & Bush, S. (2005). An evaluation of the FRIENDS programme: A cognitive behaviour therapy intervention to promote emotional resilience. *Archives of Disease in Childhood, 90,* 1016-1019.

* Stallard, P., Simpson, N., Anderson, S., & Goddard, M. (2008). The FRIENDS emotional health prevention programme: 12 month follow-up of a universal UK school based trial. *European Child and Adolescent Psychiatry, 17*, 283-289.

* Stallard, P., Simpson, N., Anderson, S., Hibbert, S., & Osborn, C. (2007). The FRIENDS emotional health programme: Initial findings from a school-based project. *Child and Adolescent Mental Health, 12,* 32-37.

Stehouwer, R. S. (1985). Beck Depression Inventory. In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques,* (Vol. 2, pp. 83-87). Kansas City, MO: Test Corporation of America.

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American Statistical Association, 54*, 30-34.

Stewart, R. E., & Chambless, D.L. (2009). Cognitive-behavioral therapy for adult anxiety disorders in clinical practice: A meta-analysis of effectiveness studies. *Journal of Consulting and Clinical Psychology, 77*, 595-606.

* Stice, E., Rohde, P., Seeley, J. R., & Gau, J. M. (2008). Brief cognitive-behavioral

depression prevention program for high-risk adolescents outperforms two alternative interventions: A randomized efficacy trial. *Journal of Consulting and Clinical Psychology, 76,* 595-606.

Storch, E. A. & Crisp, H. L. (2004). Taking it to the schools — transporting empirically supported treatments for childhood psychopathology to the school setting. *Clinical Child and Family Psychology Review, 7,* 191-193.

Strauss, C. C., Frame, C. L., & Forehand, R. (1987). Psychosocial impairment associated with anxiety in children. *Journal of Clinical Child Psychology, 16,* 235-239.

The TADS Team (2007). The treatment for adolescents with depression study (TADS): Long-term effectiveness and safety outcomes. *Archives of General Psychiatry, 64,* 1132-1144.

Timbremont, B., Braet, C., & Dreessen, L. (2004). Assessing depression in youth: Relation between the Children's Depression Inventory and a structured interview. *Journal of Clinical Child and Adolescent Psychology, 33,* 149-157.

Tolan, P. H. & Dodge, K. A. (2005). Children's mental health as a primary care and concern. *American Psychologist, 60,* 601-614.

United States Congress, Office of Technology Assessment. (1991). *Adolescent health* OTA-H-33). Washington, DC: US Government Printing Office.

United States Congress, Office of Technology Assessment. (1986). *Children's mental health: Problems and services*. OTA-H-33). Washington, DC: US Government Printing Office.

U.S. Public Health Service (2000). *Report on the surgeon general's conference on*

children's mental health: A national action agenda. U.S. Government Printing Office, Washington, DC.

Walkup, J. T., Albano, A. M., Piacentini, J., Birmaher, B., Compton, S. N., Sherrill, J. T. et al. (2008). Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *The New England Journal Of Medicine, 359*, 2753-2766.

* Warner, C. M., Fisher, P. H., Shrout, P. E., Rathor, S., & Klein, R. G. (2007). Treating adolescents with social anxiety disorder in school: An attention control trial. *Journal of Child Psychology and Psychiatry, 48,* 676-686.

Weist, M. D., Evans, S. W., & Lever, N. A. (2003). Introduction: Advancing mental health practice and research in schools. In M.D. Weist, S.W. Evans, & N.A. Lever (Eds.), *Handbook of school mental health: Advancing practice and research* (pp. 1-7). New York, US: Kluwer Academic/Plenum Publishers.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*, 688-701.

Weisz, J. R., & Jensen, A. L. (2001). Child and adolescent psychotherapy in research and practice contexts. Review of the evidence and suggestions for improving the field. *European Child and Adolescent Psychiatry, 10*, I/12–I/18.

Wicks-Nelson, R. & Israel, A. (2009). *Abnormal Child and Adolescent Psychology* (7th ed.). New Jersey: Pearson Education.

Wilkinson, L., & the APA Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594-604.

Wood, J. (2006). Effect of anxiety reduction on children's school performance and social adjustment. *Developmental Psychology, 42,* 345-349.

Woodward, L. J. & Fergusson, D. M. (2001). Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 1086-1093.

* Yu, D. L. & Seligman, M. E. P. (2002). Preventing depressive symptoms in Chinese children. *Prevention & Treatment, 5*.

Table 1.

Summary of Descriptive Characteristics and Effect Sizes for Anxiety Intervention Studies

Study	Country	Intervention Classification	Sample Size (N)	Mean Age (years)	Number of Sessions	Pre-Post SMG Effect Size (*g*)
Aune & Stiles, 2009	Norway	Universal Prevention	1439	12.6	3	N/A
Barrett, Lock, & Farrell, 2005	Australia	Universal Prevention	668	—	10	I = 0.394 C = 0.682
Barrett, Moore, & Sonderegger, 2000	Australia	Selective Prevention	20	16.3	10	I = 0.725 C = -0.303
Barrett, Sonderegger, & Sonderegger, 2001	Australia	Selective Prevention	204	12.42	10	I = 0.448 C = -0.015
Barrett, Sonderegger, & Xenos, 2003	Australia	Selective Prevention	320	12.65	10	I = 0.711 C = 0.078
Barrett & Turner, 2001	Australia	Universal Prevention	489	10.75	10	I = 0.505 C = 0.312
Bernstein, Layne, Egan, & Tennison, 2005	USA	Early Intervention	61	9.00	9	I = 0.098 C = 0.199
Bokhorst, Goossens, & de Ruyter, 1995	Netherlands	Universal Prevention	96	11.58	12	I = 0.110 C = 0.208
Calear, Christensen, Mackinnon, Griffiths, & O'Kearney, 2009	Australia	Universal Prevention	1372	14.34	5	I = 0.152 C = -0.010

Table 1, continued

Cooley, Boyd, & Grados, 2004	USA	Targeted Prevention	10	10.30	11	I = 1.399 C = N/A
Dadds, Spence, Holland, Barrett, & Laurens, 1997	Australia	Early Intervention	128	9.40	10	I = 0.762 C = 0.884
Gallegos, 2008	Mexico	Universal Prevention	972	9.89	10	I = 0.187 C = 0.223
Ginsburg & Drake, 2002	USA	Treatment	9	15.60	10	I = 0.351 C = 0.092
Hains & Szyjakowski, 1990	USA	Universal Prevention	21	16.5	9	I = 0.516 C = 0.172
Hains, 1992	USA	Universal Prevention	17	15.5	9	I = 0.787 C = -0.295
Hains, 1994	USA	Universal Prevention	19	—	8	I = 0.171 C = -0.178
Hunt et al., 2009	Australia	Targeted Prevention	136	12.05	10	I = 0.573 C = N/A
Kiselica et al. 1994	USA	Indicated Prevention	48	15.0	8	I = 1.083 C = 0.120
Lock & Barrett, 2003	Australia	Universal Prevention	737	—	10	I = 0.252 C = 0.190
Lowry-Webster, Barrett, & Dadds, 2001	Australia	Universal Prevention	531	—	10	I = 0.536 C = 0.349
Masia, Klein, Storch, & Corda, 2001	USA	Treatment	6	15.20	16	I = 1.113 C = N/A

Table 1, continued

Masia-Warner, Klein, Dent, Fisher, Alvir, Albano, Guardino, 2005	USA	Treatment	35	14.80	14	I = 1.152 C = 0.271
Mifsud & Rapee, 2005	Australia	Targeted Prevention	91	9.50	8	I = 0.364 C = -0.110
Mostert & Loxton, 2008	South Africa	Selective Prevention	46	12.50	10	I = 0.285 C = 0.163
Muris, Mayer, Bartelds, Tierney, & Bogie, 2001	Netherlands	Treatment	36	9.90	12	I = 1.127 C = N/A
Rose et al., 2009	Canada	Universal Prevention	52	10.00	8	I = 0.288 C = 0.050
Stallard, Simpson, Anderson, Carter, Osborn, & Bush, 2005	United Kingdom	Universal Prevention	197	10.00	10	I = 0.300 C = N/A
Stallard, Simpson, Anderson, Hibbert, & Osborn, 2007	United Kingdom	Universal Prevention	89	10.00	10	N/A
Warner, Fisher, Shrout, Rathor, & Klein, 2007	USA	Treatment	36	15.10	18	I = 1.766 C = 0.599
Total						I = 0.502 C = 0.216

Note: A "—" indicates that the data was not reported or could not be obtained through communication with the authors; I = Intervention Group; C = Control Group; N/A = no control group in study or immediate postintervention data was not provided.

Table 2.

Summary of Descriptive Characteristics and Effect Sizes for Depression Intervention Studies

Study	**Country**	**Intervention Classification**	**Sample Size (N)**	**Mean Age (years)**	**Number of Sessions**	**Pre-Post SMG Effect Size (*g*)**
Calear et al., 2009	Australia	Universal Prevention	1372	14.34	5	I = 0.055 C = -0.057
Cardemil, Reivich, & Seligman, 2002 [Study 1]	USA	Selective Prevention	49	11.34	12	I = 0.642 C = 0.031
Cardemil, Reivich, & Seligman, 2002 [Study 2]	USA	Selective Prevention	113	10.94	12	I = 0.418 C = 0.365
Chaplin, et al., 2006	USA	Universal Prevention	103	12.16	12	I = 0.367 C = -0.046
Clarke et al, 1993	USA	Universal Prevention	300	15.20	5	I = 0.114 C = 0.124
Curtis, 1992	USA	Treatment	19	15.80	12	I = 1.543 C = 0.312
Ettelson, 2002	USA	Early Intervention	20	—	16	I = 0.815 C = -0.525
Gallegos, 2008	Mexico	Universal Prevention	970	9.89	10	I = 0.195 C = -0.066
Gillham et al., 1994	USA	Universal Prevention	94	—	12	I = -0.035 C = -0.274
Gillham et al., 2006	USA	Indicated Prevention	44	—	8	I = 0.252 C = 0.204
Gillham et al., 2007	USA	Universal Prevention	697	12.13	12	I = 0.184 C = 0.145

Table 2, continued

Hains & Szyjakowski, 1990	USA	Universal Prevention	21	16.5	9	I = 0.989 C = 0.392
Hains, 1992	USA	Universal Prevention	17	15.5	9	I = 0.563 C = -0.057
Hains, 1994	USA	Universal Prevention	19	—	8	I = 0.571 C = 0.215
Hannan & Rapee, 1995	Australia	Indicated Prevention	19	10.90	8	I = 0.282 C = N/A
Harnett & Dadds, 2004	Australia	Universal Prevention	208	13.58	11	I = 0.003 C = -0.040
Horowitz et al., 2007	USA	Universal Prevention	281	14.43	8	I = -0.001 C = -0.173
Jaycox et al., 1994	USA	Targeted Prevention	137	11.40	12	I = 0.349 C = 0.137
Kahn et al., 1990	USA	Treatment	34	—	12	I = 2.238 C = 0.315
Kowalenko et al., 2005	Australia	Targeted Prevention	82	14.50	8	I = 0.378 C = 0.013
Liddle & Spence, 1990	Australia	Treatment	31	9.20	8	I = 1.103 C = 0.606
Merry et al., 2004	New Zealand	Universal Prevention	331	14.20	11	I = 0.204 C = 0.025
Muris et al., 2001	Belgium	Treatment	8	—	11	I = 1.578 C = N/A
O'Kearney et al., 2006	Australia	Universal Prevention	59	16.00	5	I = 0.265 C = 0.008
O'Kearney et al., 2009	Australia	Selective Prevention	157	16.00	5	I = 0.230 C = 0.118

Table 2, continued

Pattison & Lynd-Stevenson, 2001	Australia	Universal Prevention	66	10.44	11	I = -0.093 C = -0.053
Pössel et al., 2004	Germany	Universal Prevention	252	13.97	10	I = 0.080 C = 0.247
Pössel et al., 2008	Germany	Universal Prevention	301	13.68	10	I = -0.131 C = -0.240
Quayle et al., 2001	Australia	Universal Prevention	47	12.00	7	I = 0.097 C = 0.559
Ralph & Nicholson, 1995	Australia	Treatment	17	15.0	16	I = 0.159 C = 0.092
Reivich, 1996	USA	Universal Prevention	102	12.70	12	I = 0.305 C = 0.213
Roberts et al., 2003	Australia	Treatment	179	11.89	12	I = 0.208 C = 0.069
Rooney et al., 2006	Australia	Selective Prevention	117	9.08	8	I = 0.394 C = -0.121
Sheffield et al., 2006	Australia	Universal Prevention & Indicated Prevention & Combined	1360	14.34	8 8 16	I = 0.191 C = 0.191 I = 0.525 C = 0.684 I = 0.560 C = 0.684
Shochet et al., 2001	Australia	Universal Prevention	242	13.49	11	I = 0.398 C = 0.126
Shochet et al., 2002	Australia	Universal Prevention	982	13.98	10	I = 0.118 C = 0.072

Table 2, continued

Spence et al., 2003	Australia	Universal Prevention	1234	12.87	8	I = 0.299 C = -0.071
Stice et al., 2008	USA	Indicated Prevention	261	15.60	6	I = 0.949 C = 0.431
Yu & Seligman, 2002	China	Indicated Prevention	220	11.80	10	I = 0.410 C = 0.072
TOTAL						I = 0.302 C = 0.083

Note: A "—" indicates that the data was not reported or could not be obtained through communication with the authors; I = Intervention Group; C = Control Group; N/A = no control group in study or immediate postintervention data was not provided.

Table 3. *Stem-and-leaf plot of effect sizes for intervention groups in anxiety studies*

Stem	Leaf
1.8	
1.7	66
1.6	
1.5	
1.4	
1.3	99
1.2	
1.1	13, 27, 52
1.0	83
.9	
.8	
.7	11, 25, 62, 87
.6	
.5	05, 36, 16, 73
.4	48,
.3	00, 51, 64, 94
.2	52, 85, 88
.1	10, 52, 71, 87
.0	98,
-.0	

Note: Values in left column represent the ones and tenths place of the effect size, the standardized mean gain (*g*), and the values in the right column represent the hundredths and thousandths places of the effect size.

Table 4. *Stem-and-leaf plot of effect sizes for control groups in anxiety studies*

Stem	Leaf
.9	
.8	84
.7	
.6	82
.5	99
.4	
.3	12, 49
.2	08, 23, 71
.1	20, 63, 72, 90, 99
.0	50, 78, 92
-.0	10, 15
-.1	10, 78
-.2	95
-.3	03
-.4	

Note: Values in left column represent the ones and tenths place of the effect size, the standardized mean gain (*g*), and the values in the right column represent the hundredths and thousandths places of the effect size.

Table 5. *Stem-and-leaf plot of effect sizes for intervention groups in depression studies*

Stem	Leaf
2.3	
2.2	38
2.1	
2.0	
1.9	
1.8	
1.7	
1.6	
1.5	43, 78
1.4	
1.3	
1.2	
1.1	03
1.0	
.9	49, 89
.8	15
.7	
.6	42
.5	25, 63, 71
.4	10, 18
.3	02, 49, 67, 78, 94, 98
.2	04, 08, 30, 52, 65, 82, 99
.1	14, 17, 52, 84, 95
.0	03, 55, 80, 97
-.0	01, 35, 93
-.1	31
-.2	

Note: Values in left column represent the ones and tenths place of the effect size, the standardized mean gain (*d*), and the values in the right column represent the hundredths and thousandths places of the effect size.

Table 6. *Stem-and-leaf plot of effect sizes for control groups in depression studies*

Stem	Leaf
.9	
.8	
.7	
.6	06, 82
.5	59
.4	31
.3	12, 15, 65, 92
.2	04, 12, 15, 47
.1	18, 24, 26, 37, 45
.0	08, 13, 25, 31, 69, 72, 72
-.0	40, 46, 53, 57, 57, 66, 71, 87
-.1	21, 73
-.2	40, 74
-.3	
-.4	
-.5	25
-.6	

Note: Values in left column represent the ones and tenths place of the effect size, the standardized mean gain (*d*), and the values in the right column represent the hundredths and thousandths places of the effect size.

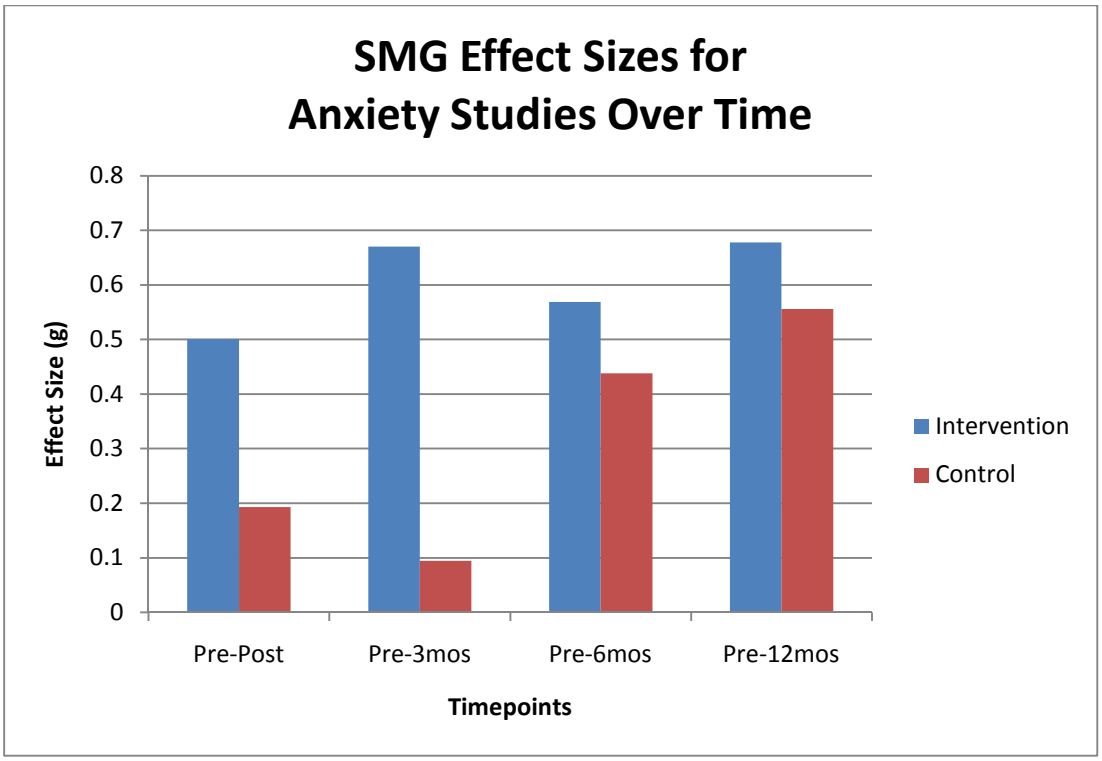

Figure 1. *Effect sizes for anxiety studies over time.*

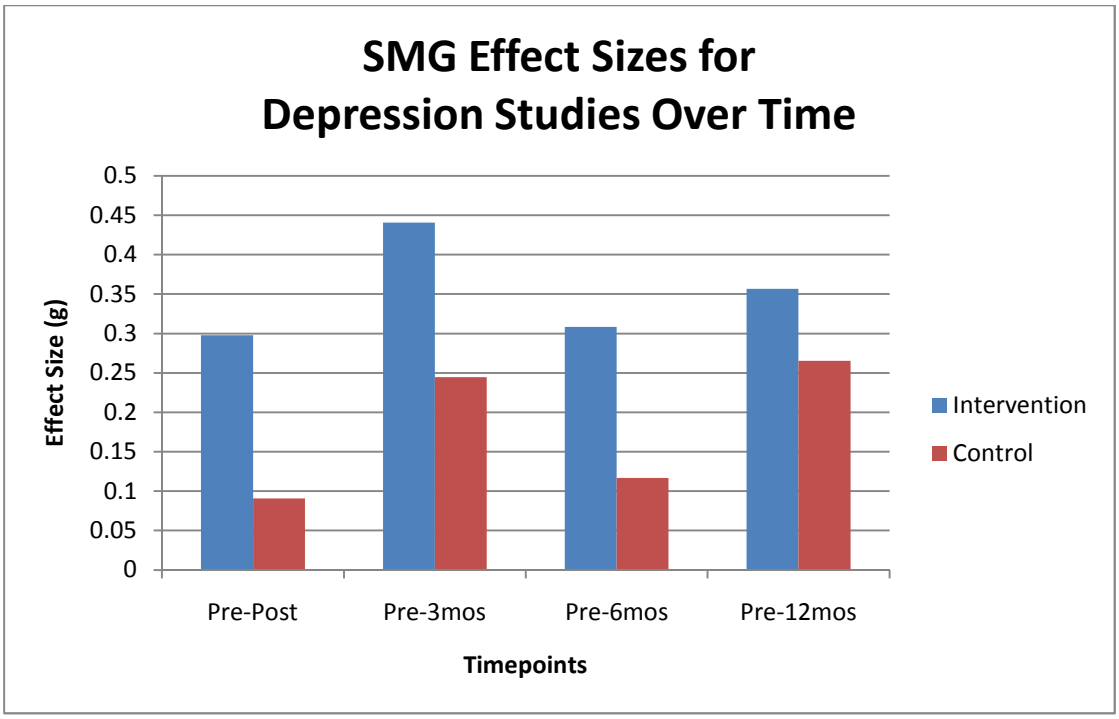

Figure 2. *Effect sizes for depression studies over time.*

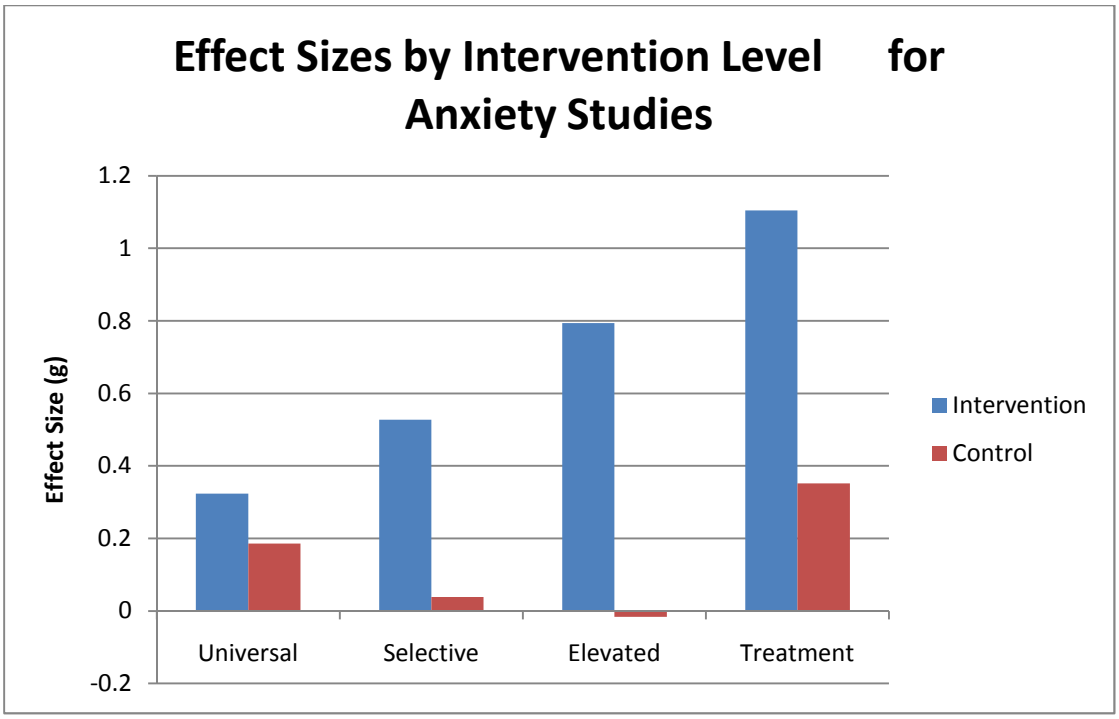

Figure 3. *Effect sizes across intervention levels for anxiety studies*

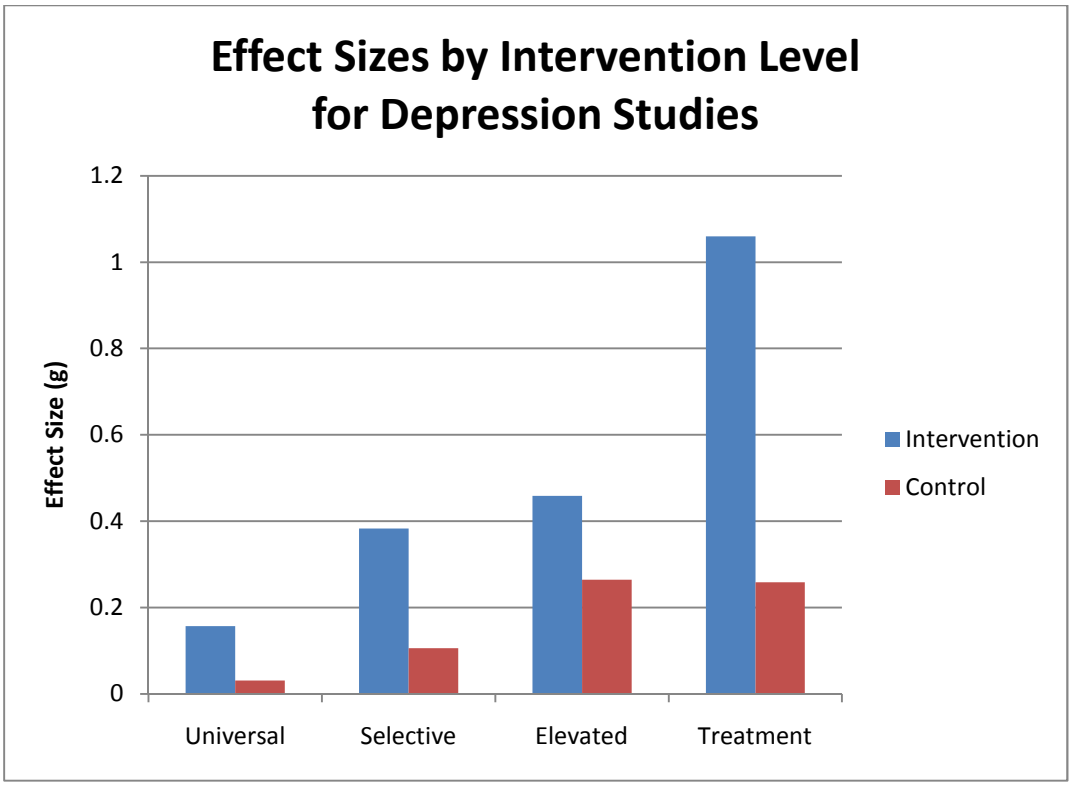

Figure 4. *Effect sizes across intervention level for depression studies*