Are treatments effective for youth with anxiety disorders? A meta-analysis

A thesis presented in partial fulfilment of the requirements for the degree

of

Master of Arts

in

Psychology

at Massey University, Wellington,

New Zealand.

Kirsten McMurray

2003

Abstract

Anxiety disorders are a disabling and enduring problem for a significant number of children and adolescents. Research in New Zealand has indicated anxiety disorders in adolescence are associated with anxiety, depression, substance dependence and academic underachievement later in life (Woodward & Fergusson, 2001). The purpose of this study was to examine the effectiveness of cognitive-behavioural treatments for childhood anxiety disorders. An additional aim was to establish the clinical significance of treatment outcomes. The design used was a meta-analysis. Seventeen outcome studies were identified for inclusion dating back to 1970, with a sample of 984 child participants, all diagnosed with a primary anxiety disorder. Though not one of the inclusion criteria, all studies employed behavioural or cognitive-behavioural approaches to treatment. Results indicated treatments were successful in alleviating anxiety symptoms and disorders as well as related areas of functioning. Overall findings indicated the average treated child was 79% more improved than untreated children or those treated with attention placebo conditions. Results also indicate that treatment effects are maintained across studies at various follow-up intervals. Moderator analyses suggested treatments to be effective for a range of anxiety disorders. Methodological limitations are identified and suggestions are made regarding future research. The relevance and application of the findings to the New Zealand context are considered.

Acknowledgments

Thank Dr Kevin Ronan for sharing his expertise and knowledge and in supervising me in completing this project. I also wish to thank John Bushnell for his helpful suggestions and comments throughout the writing and planning of this thesis.

Table of Contents

List of Figures

List of Tables

1 Introduction

Fears and anxieties are common in children and adolescents. For most children and adolescent's fears and anxieties are short-lived. Fears and anxieties that are transient and occur at a developmentally appropriate time do not present a problem (Dadds, Barrett, & Cobham, 1998; Ollendick & Ollendick, 1997). Furthermore, they are adaptive and are a normal part of childhood and adolescence. On the other hand, some children and adolescents experience maladaptive fears and anxiety resulting in considerable distress and impairment. Therefore, it is important to distinguish between normal childhood fears and anxieties and those that are excessive, persistent and maladaptive (Ollendick, 1979).

In fact, as discussed at more length in a later section, research suggests that anxiety disorders are the most commonly diagnosed group of psychological disorders diagnosed in children and adolescents (Benjamin, Costello, & Warren, 1990; Kashani & Orvaschel, 1990; McGee, Feehan, Williams, & Anderson, 1992). Negative effects of anxiety disorders include school refusal, poor academic performance, poor psychosocial adjustment, developmental delays, and difficulties in peer and family relationships (Manassis, 2000; Sanders, Dadds, & Barrett, 1995). Children and adolescents with anxiety tend to compare poorly to peers with regard to self-image, social skills, problem-solving skills, popularity and rates of interaction with peers (Dadds et al., 1998; Kashani & Orvashel). There is evidence that childhood anxiety not only causes significant distress in childhood, but is also a risk factor for poor functioning in adulthood (Pine, Cohen, Gurley, Brook, & Ma, 1998; Woodward & Fergusson, 2001).

Fortunately, there have been a number of treatment approaches developed over the past 20 years to address these problems. Many of these treatments have been evaluated in clinical trials and have indicated that this group of disorders can be successfully treated (Dadds & Barrett, 2001). This thesis focuses on those treatments.

Prior to a focus on treatment, literature on childhood anxiety disorders is first reviewed. This is followed by a review and discussion of the treatment of childhood anxiety disorders. This discussion leads into a consideration of the

means available to assess the effectiveness of treatments through a meta-analytic methodology.

1.1 When is Anxiety a Problem?

There is general agreement that children and adolescents experience fears and anxieties as part of normal development. Particular fears are expected and normal at different developmental stages. For example, young infants tend to display fears relating to loud noises or strangers. It is common for children aged 2 to 5 years to fear imaginary creatures, separation from parents and the dark. Older children and adolescents tend to have fears relating to health, school, social situations and future relationships (Dadds & Barrett, 2001; Ollendick, 1979). Evidence from research also suggests children and adolescents may have a fairly high numbers of fears (e.g., more than 10) without having an anxiety disorder (Bell-Dolan, Last, & Strauss, 1990).

Normal fears and anxieties differ from anxiety disorders in terms of the severity and persistence of symptoms. Everyone experiences fears or anxieties at some time. These feelings usually resolve when the threat is removed or some adjustment is made so the situation is no longer viewed as threatening. Fear and anxiety tend to become problematic when symptoms continue after there is no real threat (National Health Committee [NHC], 1998).

Fear and anxiety themselves are related constructs. However, a distinction is often made. Fear has been defined as a response to a real or imagined threat; anxiety, a response which lacks an obvious threat (Dadds et al., 1998; Ollendick, 1979). While some have suggested that anxiety is a more severe and complex construct (e.g., Lang, 1969; cited in Dadds et al), the *Diagnostic and Statistical Manual of Mental Disorders* ([DSM-IV], American Psychiatric Association [APA], 1994) describes anxiety and fear as manifesting similar physiological, behavioural and cognitive symptoms. DSM-IV clusters disorders with these features as Anxiety Disorders.

The DSM-IV is the main diagnostic system used to identify and diagnose anxiety disorders in children and adolescents children (APA, 1994). Diagnoses for children are consistent with those reported in adults and include: separation anxiety disorder (SAD), generalised anxiety disorder (GAD), panic disorder, specific phobias, social phobia, obsessive compulsive disorder (OCD), acute stress disorder (ASD) and posttraumatic stress disorder (PTSD). DSM-IV criteria

also note that children may also differ from adults with anxiety disorders. For example, it is characteristic for children to fail to recognise their anxiety as excessive or have difficulty recognising the cause of anxieties. Children may also express anxiety in disorganised or agitated behaviours (APA).

Overanxious and avoidant disorders, which appeared in previous editions of the DSM, are now incorporated within the Anxiety Disorders (GAD and social phobia, respectively). Separation anxiety disorder remains the only anxiety disorder specific to childhood (APA, 1994). Debate continues as to whether selective mutism should be considered an anxiety disorder of childhood or remain, as in DSM-IV, classified as an 'Other Disorder of Infancy, Childhood, and Adolescence (see Anstendig, 1999). Selective mutism has been hypothesised to be a form of social phobia, separation anxiety, and posttraumatic stress disorder (Anstendig; Kaplan & Sadock, 1998).

A defining feature of anxiety disorders is avoidance behaviour (Bernstein & Kinlan, 1997). A child or adolescent experiencing anxiety may display a number of the following symptoms: school refusal, clinginess, social isolation, and excessive shyness. While anxiety-based school refusal is not a disorder in itself, it is thought to stem from a number of diagnosable anxiety disorders, and tends to be associated with separation anxiety or a phobia (i.e. school phobia) (Last, Hansen, & Franco, 1998). Other symptoms of anxiety may include: frequent worry about the safety of self and parents, disturbed eating patterns, being easily startled or distressed, persistent nightmares, numerous stomach pains, headaches and other physical complaints, and difficulty finishing school work due to starting again or correcting 'mistakes'. Symptoms more common in young children include crying or tantrums when separated from parents and secondary enuresis and encopresis. By contrast, adolescents are more likely to exhibit comorbid depression, substance abuse, truancy, risk-taking behaviour, poor academic performance, and difficulty with social relationships (NHC, 1998).

When considering whether fears and anxieties in children and adolescents have reached an abnormal level, it is important to consider symptoms not only in terms of severity and persistence but also in terms of what is developmentally appropriate. For example, separation anxiety is common in young children beginning school. However, if separation anxiety lasts into adolescence, it is likely to be a problem (Dadds & Barrett, 2001; NHC, 1998).

1.2 Prevalence and Course

Epidemiological studies suggest that anxiety disorders are the most commonly diagnosed group of disorders in children and adolescents. Estimates of affected school aged children meeting diagnostic criteria range from approximately 5% to 20% (Bell-Dolan et al., 1990; Bernstein, Borchardt, & Perwien, 1996). For example, in a study specifically designed to investigate anxiety disorders, 21% of a community sample of 210 children aged 8, 12, and 17 years were identified as having an anxiety disorder (Kashani & Orvaschel, 1990). Benjamin and colleagues (1990) found in a sample of 300 paediatric primary care patients a 1-year prevalence rate of 15.4%. New Zealand findings from the New Zealand Dunedin longitudinal study, report prevalence rates of 3.7% at age 11 years and 7.9% at age 15 years for a community sample of 750 youth (McGee et al., 1992). In all of these studies, anxiety disorders were the most commonly reported group of psychological disorders across all age groups.

The most common anxiety disorders diagnosed in childhood are, GAD, SAD, social phobia and specific phobia (Kashani & Orvaschel, 1990; Muris, 2002). In a sample of 11-year-olds Anderson, Williams, McGee, and Silva (1987) reported 3.5% prevalence rates for SAD; 2.9%, overanxious disorder (GAD); 2.4%, specific phobia; and 1.0%, social phobia. There have been fewer studies of the less commonly diagnosed anxiety disorders, including OCD, ASD and PTSD.

Despite fewer studies, prevalence rates for OCD have been estimated at between 1% and 4% for children and adolescents (Waters, Barrett & March, 2001; Wewetzer, Jans, Müller, Neudröfl, Bücherl, Remschmidt, et al., 2001). Fewer studies yet have estimated the incidence of PTSD in child and adolescent populations. However, available findings suggests rates may be particularly high for children experiencing sexual and physical abuse, and in other populations or populations experiencing social upheaval (Anderson, 1994; Boyd, Kostanski, Gullone, Ollendick, & Shek, 2000; Pfefferbaum, 1997). One study of a community sample of over 300 youth reported a lifetime prevalence of 6% by age 18 years (Pfefferbaum). It has been hypothesised that the incidence of PTSD may be particularly high in young people as they have yet to develop adequate coping mechanisms to deal with trauma. For instance, 80% of young children compared to 30% of adults have been found to develop PTSD after a burn injury (see Kaplan & Saddock, 1998).

Again, because these disorders have until more recently been regarded as rare, they have seldom been included in epidemiological studies. Consequently, the exclusion of OCD, ASD and PTSD may have resulted in an underestimation of the prevalence of childhood anxiety disorders (NHC, 1998).

1.3 Chronicity

There is increasing evidence that for some children anxiety disorders persist over time (Anderson, 1994; Kendall, Kortlander, Chansky, & Brady, 1992; Ollendick & King, 1994). Lifetime estimates suggest up to half of those diagnosed with an anxiety disorder in childhood or adolescence will be significantly affected for at least 8 years (Keller, Lavori, Wunder, Beardslee, Scwartz, & Roth, 1992). Children who are not referred for treatment are also more likely to have long-term difficulties with anxiety and be significantly more impaired (Bernstein et al., 1996; Essau, Conradt, & Petermann, 2002). Last and colleagues (1997) found that untreated children were at risk of experiencing new disorders over time, principally other anxiety disorders (16% of the anxiety disordered sample went on to develop another anxiety disorder).

Chronicity tends to be associated with a range of problems. These include dependency, poor problem solving and social difficulties (Dadds & Barrett, 2001). Compared to non-anxious peers, Kashani and Orvashel (1990) found that anxious 12-year-olds had more problems in school, poorer self-image, and poorer interactions with peers. Anxious 17-year-olds also had these difficulties, and were more likely to have externalising problems, mood problems, and somatic symptoms compared to participants without anxiety.

Adults with anxiety disorders often report the onset of symptoms in childhood and adolescence (Dadds et al., 1998; Ollendick & King, 1998). While the link between childhood and adult anxiety is not as yet fully established, longitudinal evidence suggests clearly that childhood anxiety creates a risk for anxiety and depression in adolescence and adulthood (Dadds & Barret, 2001; Woodward & Fergusson, 2001). Childhood anxiety is also associated with an increased risk for later substance use and suicide attempts (Dadds & Barrett, 2001).

Other risk factors for anxiety disorders have also been identified and include comorbidity, age, gender, family factors, and cognition. These will now be

discussed (Anderson, 1994; Ollendick, 1979; Sanders et al., 1995).

1.4 Comorbidity of Disorders

It is generally accepted that over half of anxious youth presenting for treatment will have a comorbid disorder (Ollendick & Ollendick, 1997). Comorbid disorders include other anxiety disorders, depression, ADHD, and conduct or oppositional disorder (Anderson, 1994; Sanders et al., 1995). Comorbidity also appears to be more likely in adolescence (Anderson, 1994; Bernstein & Kinlan, 1997; Kendall et al., 1992).

The most common disorder found to be comorbid with anxiety disorders in children and adolescents is another anxiety disorder (36%; Kashani & Orvaschel, 1990). Depression is also frequently comorbid with anxiety disorders. Axelson and Birmaher (2001) report about 10-15% of anxious youth have depression. Up to one third of children with anxiety disorder will also have ADHD (Anderson, 1994). Anxiety comorbid with ADHD is more likely in younger children as the incidence of ADHD decreases in adolescence. Anderson (1994) estimate that up to one third of anxious children may have conduct or oppositional disorder. Anxiety disorders are also found to be comorbid in children with other disorders such as Tourette's disorder and other tic disorders, epilepsy, and autism.

Kendall and colleagues (1992) reported comorbidity was found to have a negative effect on treatment outcome. However, recently published findings from Kendall, Brady and Verduin (2001) have failed to support this. They found that of 173 anxious children treated with CBT, a similar percentage (68% of noncomorbid participants and 71% of comorbid participants) significantly improved. Barrett and colleagues (2001) also found no significant difference at long-term follow-up between children with comorbid diagnoses versus those without on a range of outcomes including diagnostic status.

Findings suggest the influence of comorbidity is likely to be complex and the impact on treatment outcomes remains unclear. However, there is some evidence that children with comorbid anxiety and depression show poor functioning and high levels of social anxiety, and are more likely than anxious counterparts to have mental health problems as adults (Manassis, 2000; Manassis & Menna, 1999).

1.5 Age

Findings from epidemiological studies show that the number of fears and anxieties experienced by children decreases with age (Ollendick, 1979; Gullone, 2000). Young children (8-10 years) have also been found to have more intense fears compared to older children and adolescents (King, Ollier, Iacuone, Schuster, Bays, Gullone, et al., 1989). Furthermore a study comparing prevalence rates of anxiety and depression in normative samples from worldwide studies found this trend was apparent across many cultures (Boyd et al., 2000).

There are also trends in the types of anxiety disorders experienced by different age groups. Kashani and Orvashel (1990) reported a difference in rates of separation anxiety across the three age groups in their community sample, with 18.6% of 8 year olds, 8.6% of 12 year olds, and 11.4% of 17 year olds meeting criteria. Among a sample of 6 to 8 year old school children, anxiety disorders were found to make up 23.8% of all diagnoses, and specific phobias alone had a prevalence of 21.8%, accounting for most of the anxiety diagnoses (Dadds et al., 1998). In contrast, a large study of adolescents showed an increased incidence of disorders such as panic disorder and social phobia (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993). These findings are consistent with a shift in the content of anxiety and fears across different age groups and highlight the link between anxiety disorders and developmental tasks and cognitive development (Craske, 1997; Herbert, 1994).

Research also suggests age may also be a factor in differential treatment response. Southam-Gerow, Kendall, and Weersing (2001) found older children had less favourable treatment response amongst a sample of 135 youth completing CBT programs for children with anxiety disorders. Younger children may also benefit more than older children from treatments which include parents (Anderson, 1994; Barrett, Dadds, & Rapee, 1996).

1.6 Gender

While the picture is complex, findings suggest being female can create an increased risk of developing an anxiety disorder (Dadds et al., 1998; Herbert, 1994). Prevalence studies suggest more females than males suffer from childhood anxiety (e.g., Costello et al., 1988; Kashani & Orvashel, 1990; Lewinsohn et al., 1993; McGee et al., 1992).

McGee and colleagues (1992) reported boy to girl ratios of 0.4:1 and 0.6:1 with female predominance at ages 11 and 15 years. Kashani and Orvashel (1990) found of their sample, 21% of girls and only 4.8% of boys were diagnosed with separation anxiety. Findings are consistent with adult literature reporting lifetime prevalence of anxiety disorders of 30.5% for woman compared to 19.2% for men (Kaplan & Sadock, 1998).

The literature also indicates that the occurrence of anxiety disorders in girls and boys varies as a function of disorder and age group (Anderson, 1994; Sanders et al., 1995). While overall gender ratios for social phobia tend to be similar, girls tend develop symptoms of social phobia earlier than boys with boys more likely to develop the disorder in early adulthood (Anderson, 1994; Bernstein et al., 1996; Kashani & Orvashel, 1990). Additionally, rates for overanxious disorder have been reported as similar until adolescence when females tend to outnumber males (Dadds et al., 1998).

Being female has been associated with improved response to treatment. In a treatment outcome study for youth with specific phobias, Öst and colleagues (2001), found a higher proportion of girls than boys were clinically improved following treatment. Furthermore, Mendlowitz, Manassis, Bradley, Scapillato, Miezitis and Shaw (1999) examined the effect of cognitive-behavioural group interventions with a sample of school-aged children with anxiety disorders. They found girls tended to improve more in group format treatments compared to boys. When Barrett et al. (1996) compared individual CBT with a condition including family management, girls in the condition with a family component responded better than those girls in the individual condition. Boys did equally well in both conditions.

It remains unclear why girls and boys respond differently to treatment. Mendlowitz and colleagues (1999) suggest that girls may report less anxiety posttreatment than boys because over the course of treatment boys become more willing to acknowledge their difficulties. That is, boys may initially under report their anxiety. The fit between females and therapy has also been suggested as influencing treatment outcome. For example, therapists are more often female, and girls tend to be more sophisticated in their use of interpersonal relationships to aid self-discovery and change than do boys (Weisz, Weiss, Granger, & Morton, 1995).

1.7 Family Factors

A familial or genetic component has been suggested as a factor in the development of anxiety disorders. Although it is difficult to establish with certainty whether this transmission is due to genetic or environmental factors, there is some evidence to support the existence of a genetic transmission in the development of anxiety disorders in children and adolescents (Dadds et al., 1998; Manassis, 2000; Sanders et al., 1995).

Familial studies suggest family history of anxiety disorder or other psychopathology to be a risk factor. For example, a naturalistic study found that 66% (26 of 38) children with anxiety disorders had at least one parent with a history of major affective disorder (Keller et al., 1992). Thirty-seven percent of these children had at least one parent with a diagnosis of generalised anxiety disorder, a phobia, or panic disorder. More generally, children with an anxious parent have a higher risk of developing an anxiety disorder (Bolton, 1994).

1.7.1 Genetics and Neurobiology

Anxious children often have parents with a wide range of psychiatric problems and it is thought the genetic contribution to anxiety disorders may not be specific (Dadds et al., 1998). Cummins and Ninan (2002) suggest there is a neurobiological predisposition for the development of anxiety disorders. They present a model linking neuroanatomy and neurophysiology and the expression of fears. This predisposition may result from a combination of genetic, developmental and environmental factors (e.g., experience of trauma).

Cummins and Ninan (2002) also discuss the role of behavioural inhibition, a component of temperament. Children with higher levels of behavioural inhibition have also been found to experience a number of physiological markers, putting them at increased risk for experiencing anxiety symptoms (Bernstein et al., 1996). These include higher heart rate, decreased heart rate variation, increased muscle tension, and increased salivary cortisol levels as compared to uninhibited children (Cummins & Ninan). Furthermore, behavioural inhibition appears to act as an enduring trait, and is associated with tendencies to avoid and withdraw in new or unfamiliar situations.

1.7.2 Family Processes

In addition to genetic factors, there are a number of hypotheses speculating about the link between parental and child anxiety. This section discusses some of these processes that are implicated in the development and maintenance of childhood anxiety disorders.

Silverman, Cerny, Nelles and Burke (1988) found a tendency in children to exhibit fears and anxieties modelled by their parents. As discussed in Sanders et al. (1995), anxiety can be modelled in the family and through behaviours such as parents engaging in avoidance behaviours and emphasising threat. Research has shown that parents of anxious children more often reinforce or enhance their child's avoidance strategies compared to parents of other children (Barrett, Rapee, Dadds, & Ryan, 1996). Furthermore parents who are unable to manage their own anxiety may have difficulty in helping their children to develop this skill (Hirshfeld-Becker & Biederman, 2002).

A review by Rapee (1997) concluded that parental over-control may also be associated with childhood anxiety. For example, Krohne and Hock (1991; cited in Barrett, 1998) found that anxiety in children is associated with high rates of negative feedback and parental restriction. Related to over-control, mothers of children with anxiety disorders have been found to have lower expectations of their children's coping, compared to mothers of normal controls (Kortlander, Kendall, & Panichelli-Mindel, 1997). Low expectations may be associated with protective parenting, and maintenance of poor coping and anxiety behaviours.

Other family risk factors include chronic marital conflict, loss of a parent, divorce, sexual and physical abuse and physical illness (Bolton, 1994; NHC, 1998; Sanders et al., 1995). These events may increase the risk of anxiety, as they tend to lead to changes in family routines and relationships. This may create uncertainty or perceived threat for some children.

Socio-demographics and community variables may also be risk factors and increase family stress (NHC, 1998). For example, separation anxiety disorder is more commonly found in children from single parent and low socioeconomic households (Dadds et al., 1998). Living in urban, low-income, and violent communities may increase the risk of developing an anxiety disorder (Ginsburg & Drake, 2002). These factors may place certain minority groups at increased risk for developing anxiety disorders. When Boyd and colleagues (2000) compared

prevalence rates for anxiety across normative samples worldwide, they found significant differences in the prevalence rates across countries. Eastern European countries reported higher rates compared to Western and Asian countries. The authors suggested factors such as socioeconomic status and the impact of social and political changes as risk factors for children developing anxiety. These factors may create family and financial instability, and limit access to educational opportunities and social support (Boyd et al).

Insecure attachment has for sometime been thought to be a risk factor for the development of childhood anxiety disorders (Bernstein et al., 1996). Warren, Huston, Egeland and Sroufe (1997) found this relationship in a 6-year longitudinal study. They found that the "anxious/resistant" attachment style at age 12-years consistently predicted the occurrence of anxiety disorders in adolescent participants. The authors concluded that infants experiencing this form of attachment are at increased risk for anxiety disorders as children and adolescents. It must be said that research has also highlighted that a significant proportion of children with insecure attachment do not in fact go on to develop anxiety disorders (Bernstein et al., 1996). Thus, it is likely that a number of risk and protective factors are involved, and attachment alone, like other factors reviewed, is unlikely to predict the development of anxiety disorders.

Family factors may also have an impact on treatment outcome. Research findings support the idea that parental psychopathology and parental anxiety management are important moderators of treatment outcome (Cobham, Dadds, & Spence, 1999). A number of treatment outcome studies have looked at the role of the family in CBT interventions (e.g., Barrett et al., 1996; Kendall, 1994; Mendlowitz et al., 1999). In addition to child intervention, programs involving parents include a focus on teaching parents to manage their child's anxiety through reinforcement, planned ignoring, modelling and involvement in exposure tasks. In some programmes, parents are also taught skills to manage their own anxiety (Dadds & Barrett, 2001).

Results indicate CBT interventions including a family component produce positive outcomes consistent with child interventions. However, it is unclear as yet to what extent a family component enhances treatment efficacy in the short and long-term (Barrett et al., 1998; Silverman, Kurtines, Ginsburg, Weems, Rabian, & Serafini, 1999; Spence, Donovan, & Brechamn-Toussaint, 2000).

1.8 Cognition

Cognitive factors can increase the risk of the development of anxiety disorders. Cognitive theory proposes anxiety stems from cognitive biases (Beck, Emery, & Greenberg, 1985). Negative thoughts and negative self-talk are associated with higher levels of anxiety (Ronan Kendall, & Rowe, 1994). Biases found in children with anxiety disorders tend more often to be distortions in processing and include selective attention and memory for threatening stimuli, and biased interpretation of ambiguous stimuli (Manassis, 2000).

As with adults, the resulting cognitive style is characterised by an overestimation of threat and negative consequences and an underestimation of coping (Alfano, Beidel, & Turner, 2002; Dadds et al., 1998). While this cognitive style appears to create vulnerability, prospective research is required to substantial suggestive findings.

1.9 Assessment of Anxiety Disorders Diagnosed in Childhood

Assessment of anxiety disorders will not be focused on comprehensively in this review. However, it is useful to consider issues relevant to the assessment and diagnosis of anxiety disorders in children and adolescents.

Diagnostic interviews are the most common means of diagnosing childhood anxiety disorders and many have been developed in accordance with DSM diagnostic criteria. Structured diagnostic interviews are particularly useful due to their diagnostic reliability (Ambrosini, 2000). Other tools that can be used in an assessment include behavioural observations, self-report and self-monitoring instruments, clinician ratings, parent and teacher report, and family assessment (Ronan, 1996).

A potential problem with many measures used is their adaptation from those originally developed for adults. As research on anxiety disorders in children and adolescents has not had the same attention as that for adults, it is may be inappropriate to assume measures should have the same focus. For example, the utility of generalised anxiety disorder and social phobia criteria in diagnosing children and adolescents is yet to be assessed. Additionally, many measures report lower rates of reliability compared to adult versions (Manassis, 2000; Werry, 1994).

It is therefore recommended that multimodal assessment approaches are used in an attempt to minimise these issues (Bernstein, Shaw, Dunne, Ayres, Arnold, Benedek, et al., 1997; Dadds & Barrett, 2001; NHC, 1998). For example, including behavioural observations, self-report and parent reports. Furthermore, a comprehensive assessment should include an emphasis on functional assessment (Ronan, 1996).

Consideration of developmental factors is critical in any assessment of anxiety behaviour in children. For example, young children often have normal eye-related fears (e.g., fear of the dark) and these need to be differentiated from phobias (King & Ollendick, 1997). There are also developmental differences in the presentation of anxiety disorders. Younger children with separation anxiety may experience unrealistic worry about their parents, while adolescents more commonly present with somatic complaints (King & Ollendick).

In the following section, treatment approaches for childhood anxiety disorders will be discussed. The major treatment approaches are to be described along with a summary of outcome findings.

1.10 Treatment

1.10.1 Overview: Psychopharmacological and Psychosocial Approaches

Pharmacological treatment will not be discussed in depth here. However, the current opinion on the use of drug therapy in the treatment of childhood anxiety disorders is considered briefly. There have been few outcome studies demonstrating the efficacy of medications. Selective Serotonin Reuptake Inhibitors (SSRIs) and tricyclic antidepressants (TCAs) are commonly used to treat anxious children. However, there has not been any clear evidence to support their use or in fact establish which of a number of medications are effective (Manassis, 2000; March, 2002). It is generally agreed pharmacotherapy should never be the only intervention. Some believe it can play a role in allowing skill development by allowing some alleviation of symptoms while psychological interventions are being implemented (Bernstein et al., 1997; Manassis, 2000).

A few studies have specifically explored the efficacy of psychodynamic psychotherapy for children with anxiety disorders. A study of phobic children, comparing reciprocal inhibition, psychodynamic therapy and a waitlist control

offered some support for the use of dynamic approaches. The two treatments were found to be equally effective compared to control, and treatment effects were maintained at 1 and 2 year followed up (Hampe, Noble, Miller, & Barrett, 1973; Miller, Barrett, Hampe, & Noble, 1972). More recently, a retrospective study of children with anxiety disorders treated at the Anna Freud Centre in London suggests psychoanalysis may also be effective, particularly for younger children (Target & Fogarty, 1994).

While psychodynamic treatments have received some support from meta-analyses exploring the efficacy of psychotherapy for children, the use of dynamic approaches in the treatment of anxious children and adolescents has yet to be adequately explored. Consequently, it is best considered a promising treatment (Bernstein et al., 1997; Chambless & Ollendick, 2001). In addition, this treatment presents issues surrounding the length and intensity of treatment required (Bernstein et al., 1997; Ollendick & King, 1998).

Behavioural and cognitive-behavioural approaches have been examined more comprehensively and have repeatedly been identified in a number of reviews, as potentially efficacious treatments for children and adolescents with anxiety disorders (e.g., Casey & Berman, 1985; Chambless & Ollendick, 2001; Ollendick & King, 1998; Weisz, Weiss, Alicke, & Klortz, 1987; Weisz et al, 1995). These approaches are now discussed.

1.11 Behavioural Approaches

1.11.1 Systematic Desensitisation

Originally developed by Wolpe (1958; cited in Ollendick, 1979), systematic desensitisation is grounded in the principles of classical conditioning and uses counterconditioning techniques to control fear and anxiety. Counterconditioning involves pairing an object of fear with another stimulus. It is intended to elicit a response incompatible with a fear response (Ollendick & King, 1998). The essential components of systematic desensitisation are: initiation of the incompatible response (usually relaxation); construction of a fear hierarchy; and systematic graded pairing of items in the hierarchy with the incompatible response (King & Ollendick, 1997).

Techniques involving imaginal and *in vivo* desensitisation are commonly

used and are often included as part of a more comprehensive anxiety treatment package (e.g., Cornwall, Spence, & Schotte, 1996; Hayward, Varady, Albano, Thienemann, Henderson, & Schartzberg, 2000; Kendall, 1994; Kendall, Flannery-Schroeder, Panichelli-Mindel, Southam-Gerow, Henin, & Warman, 1997; Last, et al., 1998; Muris, Merekelbach, Holdrinet, & Sijsenaar, 1998; Öst et al., 2001). In a number of outcome studies (e.g., Millar et al., 1972; Ultee, Griffioen, and Schellelkens, 1982), systematic desensitisation techniques alone have been identified as more effective than no treatment and have been established as probably efficacious treatments (Chambless & Ollendick, 2001).

It has been suggested that imaginal desensitisation may be problematic for younger children and those who find it difficult to clearly imagine stimuli and master relaxation responses. In these cases, *in vivo* desensitisation is preferred (King & Ollendick, 1997). Another reason *in vivo* techniques may be more effective is that they involve a skill development component, where children must develop skills to successfully confront their fears (Ollendick & King, 1998). In fact, evidence suggests *in vivo* desensitisation is more effective than imaginal desensitisation (Ollendick & King, 1998). For example, Ultee et al. (1982) compared imaginal and *in vivo* desensitisation in the treatment of water-phobic children. They found *in vivo* desensitisation was superior to both imaginal desensitisation and the no treatment control condition.

Emotive imagery and eye movement desensitisation and reprocessing (EMDR) are variations of systematic desensitisation. Emotive imagery involves the development of a fear hierarchy and use of an imaginal story as the incompatible response (Cornwall et al., 1996; King & Ollendick, 1997). EMDR involves the use of rapid, lateral eye movements while the patient exposes him/herself imaginally to the feared stimuli (Muris et al., 1998). The eye movements act as the incompatible response. These treatments have not been empirically validated, though there has been one study supporting the use of emotive imagery compared to no treatment for children with darkness phobia (see Cornwall et al.). By contrast, the two studies to date on EMDR both found EMDR to be less effective than *in vivo* exposure (Muris, Merekelbach, Haaften, & Mayer, 1997; Muris et al., 1998).

1.11.2 Contingency Management

Chambless and Ollendick (2001) identified contingency management (CM) as a well-established treatment for childhood phobias. CM has been particularly successful in treating children with school refusal, social and specific phobias (Dadds et al., 1998). Based on operant conditioning principles, CM seeks to reduce fear and anxiety through manipulation of consequences relating to phobic or avoidance behaviours. CM employs techniques such as: shaping, positive reinforcement and extinction to reduce fearful or anxious behaviours and to increase coping behaviours (Dadds et al; Ollendick & King). As with systematic desensitisation, CM treatments are often combined with other techniques including parental reinforcement, contract and reward systems, hierarchies, and graded exposure. CM is often included in CBT treatments (e.g., Barrett, 1998; Barrett et al., 1996; Flannery-Schroeder & Kendall, 2000; Kendall, 1994; Shortt, Barrett, & Fox, 2001; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999).

A number of outcome studies have demonstrated the efficacy of CM (Leitenberg & Callahan, 1973; Menzies & Clarke, 1993; Sheslow, Bondy & Nelson, 1983). These studies indicate CM is more effective than both no treatment and comparison treatment conditions (e.g., verbal coping and adult live modelling) (Menzies & Clarke; Sheslow et al., 1983). CM appears to be particularly efficacious in treating younger children (aged 3 to 8 years) presenting with phobias common to this age group (e.g., water and darkness phobia). An earlier study by Obler and Terwilliger (1970) also supports CM effectiveness in treating disabled children. Thirty neurologically disabled children with simple phobias were randomly assigned to reinforced practice and non-reinforced practice treatments. At posttreatment, treated children were less phobic and more able at approach tasks involving the feared objects compared to children in the control condition.

Further support for CM comes from a recent outcome study comparing CM with an attention control and a self-control treatment condition (Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999). The sample ($N = 81$) ranged in aged from 6 to 16 years and had a primary diagnosis of simple phobia, social phobia or agoraphobia. Results indicated that CM was superior to the control condition and was as effective as the self-control condition. Furthermore, CM

produced improvements that were maintained at 12-month follow up.

1.11.3 Modelling

Modelling theory, developed by Bandura (1969; cited in Ollendick, 1979), emphasises observational learning. Observational learning, or modelling, is believed to be implicated in the development of fears and phobias. The basis of modelling techniques is the demonstration of nonfearful behaviour in an anxiety-provoking situation. Through the modelling of an adaptive response, children learn new skills. Anxiety is thought to be reduced through seeing others cope in the feared situation (Kendall & Gosch, 1994; King & Ollendick, 1997). Modelling procedures are designed to enable mastery of new behaviours including coping with anxiety (King & Ollendick, 1997).

Variations of modelling include: live modelling, participant modelling and filmed modelling. Live modelling involves someone (e.g., therapist or another child) demonstrating nonfearful behaviour in the presence of the feared object. Alternatively, filmed modelling involves nonfearful behaviour being demonstrated on film. Participant modelling involves the behaviour being modelled for the child and the child is then assisted in interacting with the feared object (King & Ollendick, 1997; Ollendick & King, 1998).

Numerous studies have confirmed the efficacy of modelling in the treatment of childhood phobias (e.g., Kornhaber & Schoeder, 1975; Lewis, 1974; Murphy & Bootzin, 1973). Filmed modelling and live modelling have been identified as probably efficacious treatments. Participant modelling is a well-established and empirically supported and is more effective than filmed and live modelling and imaginal desensitisation (Chambless & Ollendick, 2001; Ollendick & King, 1998). In a study of forty water-phobic children, filmed modelling (film of peers), participant modelling (therapist assisted), combined modelling (filmed and participant) and a control condition were compared (Lewis). Results indicated the conditions including participation were more effective in reducing avoidance behaviour, and the combined condition (participant and filmed modelling) was the most effective treatment condition. These findings were consistent with those of Blanchard (1970) who found participant modelling was more effective then either live modelling or live modelling plus information in treating adults with snake

phobia.

Like other behavioural techniques reviewed, modelling is often employed as part of a CBT treatment package (e.g., Kendall et al., 1992), and may have particular utility in group interventions where opportunities for peer modelling arise, and in treatments including parents (e.g., Barrett, 1998; Beidel, Turner, & Morris, 2000; Öst et al., 2001; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999; Spence et al., 2000).

1.12 Cognitive Self-control Training

Self-instructional training is the most commonly used cognitive technique in the treatment of childhood anxiety disorders (Dadds et al., 1998). While self-instructional training is one part of a larger CBT intervention, a few studies have focused on self-talk and self-control interventions.

Graziano and Mooney (1980) conducted a study where children with severe nighttime fears were randomly assigned to a treatment or waitlist control group. The treatment group was taught to use verbal coping/self-talk and relaxation to control their fear and anxiety. The treatment group showed significant improvement at the end of treatment and improvements were maintained 2 to 3 years later. The previously mentioned study by Silverman and colleagues (1999) compared self-control and contingency management interventions with an attention control condition. The self-control condition included procedures aimed at improving children's self-observation, self-talk, self-evaluation, and self-reward. As noted earlier, both the self-control and contingency management conditions were associated with improvement posttreatment and at follow up compared to the attention control condition.

1.13 Cognitive Behavioural Therapy

Cognitive behavioural interventions utilise a variety of techniques, including those discussed above, to change thoughts and beliefs of children with anxiety disorders. Maladaptive behaviours are believed to be maintained by maladaptive cognitions and various environmental contingencies. Changes in cognition and environmental contingencies are believed to promote changes in affect and behaviour (Kendall & Gosch, 1994).

Cognitive behavioural therapy (CBT) for children with anxiety disorders was initially developed and manualised by Kendall and colleagues. CBT attempts to teach skills enabling children to recognise their anxiety and accompanying physiological reactions, identify anxious cognitions, develop adaptive coping strategies, and engage in self-evaluation. Behavioural techniques include exposure and behavioural rehearsal, relaxation, modelling, contingency management and reinforcement (Kendall & Gosch, 1994; Ollendick & King, 1998).

Since its development, there have been a number of outcome studies supporting the use of the original programme (Kendall et al., 1992) with children with anxiety disorders. The first compared a 16 session CBT intervention with a waitlist control (Kendall, 1994). Participants, aged 9-13 years (N = 47), were randomly assigned to conditions. They presented with overanxious disorder, avoidant disorder and separation anxiety and a variety of secondary comorbid disorders. At posttreatment, participants in the CBT condition had improved significantly compared to the waitlist condition as measured on a number of self, parent, and teacher report instruments. Results also indicated 64% of CBT treated children compared to 5% of the waitlist children no longer met criteria for diagnosis (Kendall, 1994). Kendall et al (1997) conducted another trial with youths (9-13 years, N = 94) diagnosed with anxiety disorders, comparing CBT to a waitlist condition. Results were consistent with the previous study with 53% of the CBT group, compared to 6% of the waitlist group, no longer meeting diagnostic criteria. One year follow up data for this study and long-term follow up (average of 3 years) for the earlier study indicated treatment gains were maintained (Kendall et al., 1997; Kendall & Southam-Gerow, 1996).

Subsequent studies have explored the efficacy of CBT treatments including a family component. Building on the work of Kendall and colleagues, Barrett et al (1996) compared CBT with a CBT plus family management (CBT+FAM). Seventy-nine participants were randomly allocated to the two 12-week treatment conditions and a waitlist control condition. Results indicated both CBT treatments were superior to the waitlist condition. Comparisons between the two treatment conditions showed the participants in the CBT+FAM had overall made greater improvements than those receiving CBT alone, particularly for younger children and girls. The added benefit of CBT+FAM was again highlighted at 12-month follow up, where 96% of the children in the CBT+FAM

group compared to 70% in the CBT group did not meet diagnostic criteria for an anxiety disorder. Six year follow up data for this study indicated treatment gains were maintained for both groups, with 85% diagnosis free (Barrett et al, 2001). However, results also indicated that there were no long-term, significant differences between the CBT and CBT+FAM in terms of diagnostic status and symptomatology.

Barrett (1998) has also evaluated CBT with a family component conducted in a family group format. Participants were randomly assigned to CBT, CBT+FAM, and waitlist. Participants assigned to CBT received treatment in groups using a standard manualised CBT intervention, and the CBT+FAM group included an additional family component where parents worked with children on management of anxiety. Posttreatment and 12-month follow up data indicated that the two treatments were both effective (65% of those treated no longer met diagnostic criteria). The CBT+FAM treatment was found to be marginally more improved than the CBT group at posttreatment and follow up. Other studies have also demonstrated support for parental involvement (e.g., Spence et al., 2000; Shortt et al., 2001).

A number of studies have explored the efficacy of youth focused CBT conducted in groups has also been explored. In a pilot study, Hayward et al (2000) randomly assigned 35 female adolescents with social phobia to a CBT group condition or a no treatment condition. Posttreatment results indicated the CBT group had significantly reduced symptoms, and there was a significant reduction in those meeting diagnostic criteria (i.e., 45% compared to 4% of the control group). However, follow up data showed no significant difference between the two groups one-year following treatment. A recently published study suggests group CBT may also be effective in treating a heterogeneous group of anxiety disordered children and adolescents. Unlike previous studies which have targeted selected groups of anxiety disorders (e.g., GAD, SAD, SOP), Lumpkin and colleagues (2002) did not exclude other anxiety diagnoses in their multiple-baseline study. Twelve children and adolescents with a range of disorders were treated and outcomes indicated improvement at posttreatment and 1-year follow-up.

In contrast, Flannery-Schroeder and Kendall (2000) found that while treated children were significantly improved compared to those in the waitlist

condition, there were significant differences between those treated with CBT individual versus group formats, with the former showing superior outcomes. In a recent study comparing the efficacy of group and individual CBT, Manassis and colleagues (2002) concluded that children with anxiety disorders improved similarly with group and individual CBT interventions.

Studies are beginning to emerge of CBT treatments developed for specific anxiety disorders. Although there has been a lack of between group designs, and sample sizes have been small, findings have been encouraging. For example, in a study of four children with panic disorder plus agoraphobia, CBT was effective in eliminating panic attacks and reducing anxiety and depression to normative levels (Ollendick, 1995). Studies have also indicated CBT may be efficacious in the treatment of children with obsessive-compulsive disorder and post traumatic stress disorder (see Deblinger, Lippmann, & Steer, 1996; Thienemann, Martin, Cregger, Thompson, & Dyer-Friedman, 2001; Waters et al., 2001)

Cognitive behaviour therapy has also been compared to an attention control placebo condition. Last and colleagues (1998) randomly assigned 56 participants with school phobia to CBT or an education support condition. Results were interesting in that they indicated that the CBT and educational support groups both showed reduced symptoms of anxiety and depression and similar rates of participants free of anxiety disorder diagnoses following intervention. Both conditions were equally effective in returning children to school, raising questions about the superiority of CBT over educational and supportive interventions.

Ginsburg and Drake (2002) evaluated a school-based group CBT treatment for anxious African-American adolescents compared to placebo. Participants were randomly assigned to a CBT or attention control group. At posttreatment, 75% of the CBT group and 20% of the attention control group were diagnosis free. Although this study had a small sample size (n = 12), results indicated group CBT was effective for African-American adolescents and that treatment could be successfully conducted in a school setting. Beidel et al. (2000) also demonstrated the effectiveness of a CBT group intervention compared to an attention control group for children with social phobia. Following treatment, 67% of the CBT group and 5% of the attention control group were diagnosis free. Follow-up assessment at 6-months indicated treatment gains were maintained.

In summary, CBT has been identified as an efficacious treatment for children with anxiety disorders (Chambless & Ollendick, 2001). While there have been a number of studies published on the efficacy of CBT, including those reviewed here using attention placebo, the approach has yet to be established as more effective than other treatments (Ollendick & King, 1998). Additionally, some mixed findings indicate that the relative efficacy of group and family treatments requires further clarification.

1.14 Preventative Interventions

Research has produced information on risk factors associated with the development of anxiety disorders. This enables at risk groups of children to be identified and targeted. While not the focus of this current study, there is increasing interest in developing interventions aimed at preventing the development of anxiety and stress-related problems in children. As discussed earlier, anxiety disorders are the most common disorders diagnosed in children and adolescents. Evidence reviewed earlier also suggests that these disorders can be both severe and chronic. Therefore, the development of preventative approaches is considered worthwhile (Sanders et al., 1995; Dadds, Spence, Holland, Barrett, & Laurens, 1997).

Large-scale programmes have emerged as the most cost effective means of reducing the incidence of disorders in children (Sanders et al., 1995). To date the only programme of this kind specifically targeting children at risk of developing anxiety disorders has been the Queensland Early Intervention and Prevention of Anxiety Project. This school-based project screened 1786 children for anxiety problems. Based on teacher and self-reports, 128 children were identified as at risk for developing anxiety disorders (Dadds et al., 1997). Children were randomly assigned to a 10-week school-based child and parent intervention or a monitoring group. While posttreatment results indicated that both groups showed improvement, this was only maintained in the treatment group and those improvements were still present at 2-year follow up (Dadds, Holland, Laurens, Mullins, & Barrett, 1999). Results suggest early intervention programmes have potential in successfully targeting at risk youth.

1.15 The New Zealand Context

There is little research specific to the treatment of childhood anxiety disorders in New Zealand. Epidemiological research (McGee et al., 1992) has indicated anxiety disorders to be a significant problem for New Zealand children. As stated earlier, prevalence rates of 3.7% at age 11 years and 7.9% at age 15 years were reported for a community sample of 750 New Zealand youth.

Findings from the Christchurch Health & Development study also indicate Mäori youth were 1.5 times more likely than non- Mäori youth to have a mental health disorder (Ministry of Health [MoH], 1998). Based on these findings, we could hypothesise that Mäori may be at increased risk of developing a childhood anxiety disorder.

To date there have been no published treatment outcomes studies conducted in New Zealand. The NHC (1998) guidelines discuss children as a special population. Consistent with international literature (e.g., Bernstein et al., 1997) the guidelines recommend the use of cognitive-behavioural interventions for treating children with anxiety disorders (NHC). As these treatments have been identified as efficacious in outcome studies overseas, it seem reasonable to suggest they would also be effective in a New Zealand setting. Some multiple-baseline single case research has pointed to that potential (Girling-Butcher & Ronan, 2003; Huzziff & Ronan, 2003). Based on the current review, 'best practice' would likely include family involvement (e.g., Girling-Butcher & Ronan) in a CBT focused programme along with useful adjuncts (e.g., school liaison, medication).

Information on efficacious treatments is important in the current climate where child/tamaiti and youth/rangatahi mental health has been identified as a priority area for improved service delivery and workforce development (Health Funding Authority, 2000; Mental Health Commission, 1998). Child and youth services have been found to be dramatically deficient in services and expertise. Training has therefore been identified as a priority.

It is likely there are deficits in the services offered to children with anxiety disorders in New Zealand. Ideally, the training of child mental health workers will be driven by reliable information indicating the skills required of workers and services to deliver effective interventions. It is intended and hoped that this current study will add to knowledge related to effective interventions.

1.16 This Review

Meta-analysis is one method for summarising a body a body of research. Increasingly results from meta-analyses are being used to inform decision-making about health policy and future research directions (Gilbody & Petticrew, 1999). This study used meta-analysis to evaluate the effectiveness of psychological treatments for anxiety disorders in children and adolescents.

The aims of this review were twofold. The first was to contribute to knowledge on the effectiveness of treatments for childhood anxiety disorders by conducting a meta-analysis of treatment outcome studies[1]. In addition, there was interest in the clinical significance of treatment outcomes and whether CBT treatments are more effective than other therapies.

Second, this review assessed differing methodologies for calculating effect sizes. In a meta-analysis each study is represented by an effect size. The effect size is a statistic that indicates the magnitude of an effect (e.g., overall treatment outcome) and standardises outcomes across studies so they can be combined or compared (Hunter & Schmidt, 1990; Lipsey & Wilson, 2001). The aim here was to compare two methods of calculating effect sizes. Meta-analysis theory will now be discussed followed by a review of relevant meta-analyses of child psychotherapy outcomes.

1.16.1 Meta-analysis

Glass first used the term 'meta-analysis' (1976; cited in Hunter & Schmidt, 1990). Meta-analysis is a quantitative technique for summarising the results of an existing body of research (Glass, McGaw & Smith, 1981). The purpose of a meta-analysis is to identify a set of studies, synthesise the findings from these studies quantitatively and generate an overall conclusion (Lipsey & Wilson, 2001; Rosenthal & DiMatteo, 2001).

---

[1] To date there have been no published meta-analyses of psychological treatments for childhood anxiety disorders. However, recent correspondence with a leading researcher in the area of outcome research (P. C. Kendall, personal communication, 10 February, 2003) indicates he and colleagues are working on a related meta-analysis of for treatments with differing treatment components. A review of empirical literature on treatments for children with internalising disorders has however recently been published (Compton, Burns, Egger & Robertson, 2002). This review indicated a range of treatments (i.e., behavioural and cognitive-behavioural) were effective in treating symptoms of anxiety but did not include meta-analytic analyses.

Since the 1970s, methods of meta-analysis have been developed by a number of researchers in the behavioural and social sciences (e.g., Smith and Glass, Rosenthal and Rubin, Schmidt and Hunter, Hedges and Olkin). Meta-analysis has become a widely accepted means of analysing primary data (Clarke & Oxman, 2002; Gilbody & Petticrew, 1999).

As in this review, meta-analysis is commonly used to establish the overall effect of various forms of treatment, including psychotherapy (e.g., see Casey & Berman, 1985; Smith & Glass, 1977; Weisz et al., 1987). In synthesising studies, meta-analysis addresses difficulties commonly found in social and behavioural science research where sample size and observed effects may be small (Hedges & Olkin, 1985; Hunter & Schmidt, 1990). In combining studies, meta-analyses have increased power and thus are better able to reach more firm conclusions (Lipsey & Wilson, 2001). As meta-analysis may include data from a diverse range of studies and participants, external validity is also increased (Kavale & Forness, 2000).

Additionally, if studies are found to have varying effects, the source of variation can be investigated in moderator analyses (Borenstein & Rothstein, 1999). For example, in psychotherapy meta-analyses, the impact of age or gender on treatment outcomes is often examined (e.g., Michael & Crowley, 2002; Weisz et al., 1987; 1995).

1.16.2 Criticisms of Meta-analyses

The first meta-analyses of psychotherapy outcomes were criticised for combining studies which varied widely in terms of the interventions and outcome measures used (commonly known as 'combining apples and oranges') (Eysenck, 1994; Sharpe, 1997). Researchers have since developed methods to address this problem. It is now more common for a meta-analysis to include studies on a single treatment approach or population (e.g., Durlak, Fuhrman, & Lampman, 1991; Michael & Crowley, 2002; Rector & Beck, 2001).

Meta-analyses have also been criticised for including studies with poor methodological quality (Eysenck, 1994; Hunter & Schmidt, 1990). In the case of psychotherapy outcomes, a number of meta-analyses have shown that studies with increased methodological rigor are associated with larger effect sizes (Lipsey &

Wilson, 1993; Michael & Crowley, 2002; Weiss & Weisz, 1990). Practices such as only including well-controlled studies (e.g., including randomisation, a control comparison group and standardised measures), and weighting studies based on their sample size and methodological strength address this issue.

Finally, meta-analyses may be biased in their inclusion and exclusion of studies (Eysenck, 1994; Sharpe, 1997). In the process of identifying studies, it is inevitable that some studies will be overlooked. The concern is that due to publication bias, effects may be overestimated or effects found that do not exist, as few studies with statistically nonsignificant or negative effects are likely to be included (Glass et al., 1981; Sharpe). While this is a well-founded criticism, Hunter & Schmidt (1990) have researched this issue at length, finding the results of unpublished studies are similar to those of published studies, suggesting 'missed' studies are unlikely to alter the overall result. In addition, analyses such as the calculation of a Fail-safe *N* statistic go some way to addressing this issue. This statistic estimates the number of studies supporting the null hypothesis that would be needed to reduce the mean effect of a meta-analysis to a statistically nonsignificant effect (Lipsey & Wilson, 2001).

1.16.3 Effect Sizes in Meta-analyses

As mentioned earlier, one aim of this meta-analysis was to compare methods of calculating effect sizes. The convention in meta-analysis is to use posttreatment scores as the unit of measurement in the calculation of effect sizes. Smith and Glass (1977) calculated effect sizes based on the mean difference between treatment and control subjects scores measured at posttreatment. Benchmark studies in the area of outcomes of psychotherapy with children have also followed this method (e.g., Casey & Berman, 1985; Durlak et al., 1991; Kazdin, Bass, Ayers, & Rodgers, 1990; Weisz et al., 1987; 1995). When employing this method, an effect size simply represents the distance on the dependent variable, as measured in standard deviation units, that the average subject in the treatment group is from the average subject in the comparison group following treatment (Cohen, 1988).

Effect sizes can also be calculated based on gain scores. Gain scores refer

to the mean change in score between pretreatment and posttreatment. Gain scores provide an indication of the within-group magnitude of change and have generally been discussed in relation to within-group designs (Cohen, 1988; Hunter & Schmidt, 1990). Recently, meta-analyses have begun to calculate both between and within-group effect sizes (e.g., Michael & Crowley, 2002; Rector & Beck, 2001).

One advantage of the use of within-group gain score effect size methodology is that it allows for comparison and inclusion of studies that do not have a comparison group (Michael & Crowley, 2002). Often these are effectiveness studies conducted in naturalistic settings, versus analogue or efficacy studies predominant in reviews of psychotherapy outcomes. Another advantage of calculating within-group gain score effect sizes is their utility in evaluating the impact of treatment on variability and estimating the change in control groups from pretreatment to posttreatment that between-group effect sizes can not take into account (Michael & Crowley, 2002; Weiss & Weisz, 1990).

There are problems associated with within-group effect sizes. First, the benefits of randomisation are lost as each group becomes the equivalent of a before-and-after study (Clark & Oxman, 2002; Glass et al., 1981). Effect sizes that lack a comparison with a control group have also been shown to produce inflated effect sizes (Michael & Crowley, 2002; Rector & Beck, 2002, Western & Morrison, 2001). For example, Lipsey and Wilson (1993) found pretreatment to posttreatment effect sizes were 61% larger than those derived from treatment and control group comparisons.

Western and Morrison (2001) warn that within-group effects should be interpreted with caution as they are confounded with a number of threats to validity which can not be assessed due to the lack of comparison with a control condition. Conversely, Cohen (1988) and Hunter and Schmidt (1990) have suggested that within-group designs can be superior to between-group designs. Particularly when measures are reliable and sample sizes are small, the within-group design will have higher power than the between-group design. Within-group designs also have the potential to present data on individual treatment effects.

A further alternative is to calculate an effect size based on the difference in gain scores for treatment and control groups (e.g., comparison of pretreatment to

posttreatment change outcome scores). This method produces a between-group effect size derived from standardised difference (gain) scores rather than standardised units of the dependent variable (Lipsey & Wilson, 2001). Analyses of standardised mean differences based on gain scores are carried out in the same manner as analyses based on posttreatment scores (Clarke & Oxman, 2002). This method addresses the problems of within-group designs, including loss of the benefits of randomisation and lack of comparison with a control.

Gain scores may have a number of advantages over posttreatment scores. For example, the distribution of gain scores may be less skewed than post-treatment outcomes and gain scores are likely to estimate mean difference with greater precision (i.e. narrower confidence intervals) (Clarke & Oxman, 2002). As the standard deviation of gain scores is often smaller than that of posttreatment scores, gain scores may also lead to larger effect sizes (Rosenthal, 1994).

The inclusion of pre-treatment outcomes in analyses allows for consideration of any pre-treatment differences between groups. The omission of pretreatment outcomes has been considered a limitation of posttreatment effect size only calculations. Although it should not be a problem in meta-analysis where differences should average out across studies, this does highlight that gain scores can deal with groups that are unequal at pretreatment, thus correcting for unlucky randomisation (Glass et al., 1981; Hunter & Schmidt, 1990). Gain scores also examine the variability in treatment effects for individual subjects which is overlooked in analyses that compare posttreatment scores only (Hunter & Schmidt, 1990). Finally, there does seem to be practical value in talking about the differences in gain between treatment and control conditions.

As a contrasting view, Glass and colleagues (1981) concluded that while there are no rules about which method is best, they recommended avoiding gain scores and using final status (e.g., posttreatment) in the calculation of effects when subjects have been randomly assigned to treatment and control groups. Posttreatment scores have also been said to have the advantage of being more clinically relevant and more interpretable to audiences (Glass et al., 1981; Duggan, Fenton, Dardennes, El-Dosoky, & Indran, 2002).

Additionally, as calculations derived from posttreatment scores have been the convention in meta-analysis of psychotherapy outcomes to date, it would seem there is considerable merit in using this method to allow for direct comparison

with previous findings. The use of posttreatment scores is often more practical as outcome studies still tend to report mean pretreatment and posttreatment scores and much less commonly report data on gain scores, making the estimation of gain score standard deviations necessary. Without such data, it becomes very difficult to estimate the variability of gain scores (e.g., sample skewedness) for between-group comparisons. Meta-analysis researchers thus become reliant on seeking data from authors (Clarke & Oxman, 2002).

It has been suggested that effect sizes derived from gain scores and posttreatment scores can be combined. There is some evidence that overall results from effects seem to be similar (P. Herbison, personal communication, 27 February, 2002; Clarke & Oxman, 2002). In instances where there are no pretreatment differences between groups, results from gain scores will simply reflect the difference between posttreatment means (Glass et al., 1981; Lipsey & Wilson, 2001). However, there may be times when a particular method of calculating effects may be preferred. The use of gain scores when groups are unequal at pretreatment is the most obvious example.

More investigation is needed to establish whether or not there is a significant difference between the two methods of calculating effects. There also continues to be disagreement as to which index is more clinically relevant (Duggan et al., 2002; Glass et al., 1981; Rector & Beck, 2001). Furthermore, there are instances where it is difficult to establish which method meta-analysis researchers have actually employed and why precisely they have chosen a particular method (i.e. whether effects were calculated from mean gain scores or from mean posttreatment scores). These issues are seldom discussed in detail in the meta-analysis literature and to this author's knowledge, there has not been a direct comparison of these methods (posttreatment versus gain score) for calculating between-group effect sizes.

1.16.4 Clinical Significance in Meta-analysis

A problem with effect sizes is that while they tell us the magnitude of change on an outcome measure, they do not tell us whether the change is clinically meaningful. A statistically significant effect does not indicate whether the treatment had a significant effect on client's lives. It has become more common

for treatment outcome studies of childhood anxiety disorders to include a measure of clinical significance. However few meta-analyses include such evaluations (Kendall, Flannery-Schroeder, & Ford, 1999). Such evaluation would appear to be an important addition to the calculation of effect sizes which do not answer questions about the clinical meaningfulness of treatment effects.

While large effect sizes are more likely to indicate clinically significant change, Rosenthal (1984; cited in Meyer & Marks, 1995) points out that small effects can also be of clinical importance. Alternatively, it is not uncommon for meta-analyses to report substantial effect sizes for psychotherapy outcomes, but subsequent comparisons with normative samples indicate participants have not moved to within normal ranges of functioning following treatment (Jacobson & Truax, 1991). In psychotherapy outcome research, clinically significant improvement can be defined as change that indicates that by the end of treatment participants are performing within the range of normally functioning people (Jacobson, Roberts, Berns, & McGlinchey, 1999). There are a number of methods for measuring clinical significance.

One approach is the calculation of normative effect sizes, comparing the scores of a treatment group posttreatment with those of a normative sample. If the resulting effect size indicates there is not a significant difference between the two groups, the treatment group is considered to have improved significantly to within a normative range (Durlak et al., 1991). Social validity can also indicate clinical significance (Kazdin, 1977; cited in Kendall, Flannery-Schroeder, et al., 1999). Social validity involves examining the acceptability of treatment goals and procedures, the meaningfulness of change as evaluated by clients and their significant others (Foster & Mash, 1999).

Another approach is the Reliable Change Index (RCI) (Jacobson & Truax, 1991). The RCI also involves comparison with a normative population and ensures the magnitude of change exceeds measurement error and thus is statistically reliable and clinically significant (Jacobson et al., 1999). Kendall and Grove (1988; cited in Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999) have also developed an approach known as normative comparison. In this approach, clinical significance is defined as "end-state functioning that falls within a normative range on important outcome measures" (p. 285, Kendall, Marrs-Garcia, et al., 1999). The normative comparison method is widely used in treatment outcome

studies for childhood anxiety disorders (e.g., Barrett et al., 2001; Flannery-Schroeder & Kendall, 2000; Kendall, 1994; 1997; Shortt et al., 2001; Silverman, Kurtines, Ginsburg, Weems, Lumpkin, et al., 1999; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999).

Recently, Sheldrick, Kendall and Heimburg (2001) combined the normative comparison and RCI approaches to assess clinical significance of treatments for conduct disordered children. In combining the two methods, it becomes possible to determine whether the magnitude of change is big enough to be considered clinically significant and establish whether treatment participants have returned to within normal limits (Sheldrick et al.). This combined method was therefore used in this meta-analysis.

1.17 Meta-analyses of Child Psychotherapy Outcomes

Smith and Glass (1977) conducted the first notable meta-analysis of psychotherapy outcomes. The results of 375 outcome studies investigating the effects of psychotherapy and counselling were included. Findings provided considerable support for the efficacy of therapy. The average client receiving therapy was found to be 75% better off than untreated controls.

To date, there have been a number of large meta-analyses investigating the effectiveness of treatments for children and adolescents with psychological or behavioural problems (Casey & Berman, 1985; Kazdin et al., 1990; Weisz et al., 1987; Weisz et al., 1995). Results were drawn from more than 300 treatment outcome studies of children and adolescents aged 4 to 18 years (Weisz & Hawley, 1998). To summarise, these four meta-analyses have generated effect sizes ranging from .71 to .88. The mean effect size across the four meta-analyses was 0.80 indicating that the 79% of the treated children were found to be better off than untreated controls. As Cohen's (1988) guidelines indicate 0.80 to be a large effect size, there is strong support for psychological treatments for children.

Meta-analyses have also highlighted the differential effects of various treatment approaches in child psychotherapy. For example, Weiss and Weisz (1995) compared the effectiveness of behavioural and nonbehavioural psychotherapies. They found behavioural treatments were associated with larger effect sizes compared to nonbehavioural treatments (0.85 compared to 0.42, respectively). Casey & Berman (1985) also found the greatest improvements were

almost always associated with a form of behavioural therapy. Overall results also suggest that treatments are more effective when they target specific problems (Weisz et al., 1995).

Of relevance to this review, Durlak et al. (1991) focused solely on the effects of cognitive-behaviour therapy with maladaptive children. Findings were positive for effect sizes across 3 age groups: 0.92 for ages 11-13 years, 0.55 for ages 7-11 years, and 0.57 for ages 5-7 years. Recently, in a review of treatments for child and adolescent depression, Michael and Crowley (2002) found an effect size of 0.72 for psychological treatments. A large majority of the treatments were behavioural or cognitive behavioural.

Findings from previous meta-analyses also suggest gender and age moderate treatment outcome. As seen, Durlak and colleagues (1991) found better treatment outcomes for older compared to younger children. Weisz and colleagues (1995) and Michael and Crowley (2002) also found adolescents did moderately better than younger children. Given the predominance of cognitive-behavioural interventions, it has been suggested that older children may benefit due to their more developed cognitive abilities (Weisz et al., 1995). Furthermore, Weisz et al. also found a significant relationship between female gender and improved treatment outcome. This pattern was also found by Casey and Berman (1985) and Michael and Crowley.

Child psychotherapy has been shown to be effective, producing similar magnitude of effects to those found for adults. Follow up data also indicates treatments are generally effective in the long-term (Weisz & Hawley, 1998). These previous studies suggest that a meta-analysis of psychological treatments for childhood anxiety disorders is likely to produce a positive outcome. Based on earlier findings it is predicted that treatments incorporating behavioural and cognitive behavioural techniques and targeting specific anxiety disorders to be associated with larger mean effect sizes.

In summary, the purpose of this meta-analysis was to:

I. Examine the effectiveness of treatments for childhood anxiety disorders.

II. Compare two methods of calculating treatment effects.

III. Establish the clinical significance of treatment effects.

2 Method

2.1 Inclusion Criteria

To be consistent with previously published meta-analyses addressing the effectiveness of psychological interventions with children, a definition of psychotherapy was used following benchmark studies in this area (e.g. Casey & Berman, 1985; Kazdin et al., 1990; Weisz et al., 1987; Weisz et al., 1995). Psychotherapy was defined as "any intervention designed to alleviate psychological distress, reduce maladaptive behaviour, or enhance adaptive behaviour through counselling, structured or unstructured interaction, a training program, or a predetermined treatment plan" (Weisz et al., 1995, p. 452). Consistent with benchmark studies, interventions were included whether carried out by trained professionals, professionals in training or paraprofessionals. Studies of pharmacological treatments, interventions intended only to increase knowledge, and those where the focus was prevention, were excluded (see also Casey & Berman, 1985; Kazdin et al., 1990; Weisz et al., 1987; 1995).

As meta-analysis literature suggests that high quality studies are associated with more substantial effect sizes, the aim was to select studies with more rigorous designs (e.g., see Green & Hall, 1984; Hunter & Schmidt, 1990; Weiss & Weisz, 1990). Studies were included in this meta-analysis based on the following inclusion criteria (see also Weisz et al., 1995): (a) interventions met the definition of psychotherapy; (b) the study was published between 1970 and 2002;[2] (c) the study sample consisted of children or adolescents diagnosed with an anxiety disorder using an established structured interview and DSM criteria; (d) the focus of the study was the treatment of anxiety disorders in children; (e) the participants were randomly assigned to groups;[3] (f) the design included a comparison control

[2] 1970 was chosen as the cut off date consistent with Durlak et al (1991) and Kazdin et al (1990). No randomised controlled trials of psychological treatments specifically for children and adolescents had been identified in previous meta-analyses which included studies published pre 1970 (e.g. Casey & Berman, 1985; Weisz et al., 1987; 1995). It was therefore unlikely that studies would be missed. A recently published meta-analysis of treatments for child and adolescent depression also suggests the majority of treatment studies in the childhood anxiety area were published after 1990 (Michael & Crowley, 2002).

[3] Outcomes of non-random design studies have been investigated (e.g., Shadish & Ragsdale, 1996; 1998). Results suggest under some circumstances they produce acceptable results compared to random design studies. However, non-random design studies are more likely to have pretreatment differences between-groups. As our aim was to establish the efficacy of treatment compared to control conditions, and as non-random studies can not be reliably contrasted to random design studies, random design studies only were included.

group;[4] (g) measures were administered at pretreatment and posttreatment; and (h) statistics were reported from which effect sizes could be calculated (means, standard deviations, sample sizes).

Studies were excluded if: (a) interventions did not meet the definition of psychotherapy; (b) the sample age ranged above 20-years-old; (c) the focus was on anxiety symptoms rather than a diagnosed anxiety disorder; (d) there had not been adequate diagnosis of an anxiety disorder (i.e., an established interview was not used); (e) anxiety was secondary to a medical condition or another disorder (e.g. a developmental disorder); (f) the design did not include a comparison with an untreated or attention control group; or (g) follow-up data only was reported and no pre and post-treatment data were available.

Single-subject and within-subject designs were also excluded. This is consistent with benchmark meta-analyses in the area. It is also in keeping with the objectives of this review, where between-group comparison was of interest. Some researchers consider their exclusion neglectful as methods are now available for calculating effect sizes from these studies (e.g. Allison & Gorman, 1993; Giles, 1990). However, this study is in line with the view that the inclusion of these studies remains problematic and effect sizes from single-subject designs are simply not comparable to group derived effect sizes and are therefore best considered separately (Hunter & Schmidt, 1990).

2.2 Search Procedure

A literature search was conducted to identify outcome research of psychological treatments for children and adolescents with anxiety disorders. Following common practice in many psychological outcome meta-analyses, only published studies were included (e.g Casey & Berman, 1985; Kazdin et al., 1990; Meyer & Mark, 1995; Westen & Morrison, 2001; Weisz et al., 1987; 1995). A number of researchers have cited concerns regarding the exclusion of unpublished studies causing a source bias (e.g. Glass et al., 1981; Rosenthal & DiMatteo, 2001). However, others (e.g. Hunter and Schmidt, 1990: Weiss & Weisz, 1990) discuss the results of meta-analytic research showing unpublished studies have not

[4] Consistent with benchmark studies, the comparison control group could be a no treatment, waitlist condition or an alternative treatment condition not thought to be efficacious and acting as an attention control. It must also have been drawn from the same population. Muris, Meesters, & Gobel (2002) was excluded on the basis that the authors stated the comparison group, Emotional Disclosure, although less optimal than the CBT treatment had been demonstrated to have potential in reducing anxiety in children.

produced effect sizes clinically or significantly different from those of published studies. Instead, they suggest any bias or difference in observed effects may be due to the methodological quality of studies, with higher quality studies more likely to be published and associated with larger effects. It was decided that maximising the quality of studies was the priority and, the exclusion of unpublished studies also avoided any selection bias due to difficulties in systematically identifying unpublished studies (Dickersin, Scherer & Lefebvre, 1994).

Search methods recommended by Hunter and Schmidt (1990) were followed to ensure a thorough review of published literature. This included: (a) use of two or more databases; (b) searching the references of previous reviews; and (c) making contact with relevant scholars for relevant studies.

Published studies were identified by searching computerised databases (MEDLINE, PsycINFO) for the period 1970 to December 2002. The following keywords were used: anxiety, anxiety disorders, phobia, fear, treatment, and therapy. Articles were limited to studies focusing on treatment or therapy, where the sample population were children and adolescents, and those published in English.

The search resulted in a total of 335 studies. Abstracts were examined and the inclusion criteria applied. A large number of the identified studies were excluded as the sample was limited to adults or was not a treatment outcome study (n = 62). For studies that were treatments of children with anxiety disorders, the majority that were excluded were those that the sample had a disorder other than an anxiety disorder as the primary diagnosis (n = 155), single case designs (n = 52), a control comparison was not included (n = 22), or the intervention included drug therapy (n = 21). A further two studies were excluded as they were based on previously published data, and two were excluded as they were prevention studies. If the remaining studies could neither be systematically included nor excluded on the basis of available information, copies of the articles were obtained for further evaluation.

Reference lists of all studies and relevant reviews were checked to identify any additional studies. Issues of the last 10 years of the Journal of Consulting and Clinical Psychology and the Journal of the American Academy of Child & Adolescent Psychiatry were manually searched, as these journals were thought

most likely to publish relevant treatment outcome studies.[5] Three key authors who actively conduct outcome research in the area of child and adolescent anxiety disorders were contacted to inquire about any additional studies that had not been identified. On the recommendation of one these authors (P. C. Kendall, personal communication, 7 February, 2003) one other researcher was contacted, resulting in four opinions being sought. These additional methods were employed as they are recommended to ensure a thorough search, and are commonly practiced in meta-analysis research (Casey & Berman, 1985; Cooper, 1998; Durlak et al., 1991; Hunter & Schmidt, 1990; Weisz et al., 1995). Two of the authors contacted responded.[6] No further studies were identified through these means.

2.3 Outcome of Search

Nineteen studies, reporting 26 comparisons between treatment and control conditions, were identified for this meta-analysis.[7] Seventeen of the 19 studies were treatment outcome studies and two were follow-up studies (Barrett et al., 2001; Kendall & Southam-Gerow, 1996). Given that these follow-up studies related to earlier studies, though they were included, the sample was considered to be $N = 17$. This produced a sample size of 984 participants who were initially randomised to treatment and control conditions and 885 (90%) participants who completed these conditions. The included studies are marked in the reference section.

2.4 Coding Procedure

Following recommendations in the meta-analysis literature (e.g. Glass et al., 1981; Lipsey & Wilson, 2001; Wolf, 1986), coding of study characteristics was inclusive and covered a wide range of variables (77 in total). Studies were coded based on variables grouped under the following categories: study context, sample characteristics, methodology, and treatment characteristics. In addition to providing the basis for descriptive and effect size statistics, variables formed the

[5] This was supported by the fact that 12 (71%) of the 17 studies ultimately included in this meta-analysis were published in these two journals.

[6] One of the authors was unable to respond due to being on official leave but was available to provide feedback at a later date (P. M. Barrett, personal communication, 13 February, 2003) and one other author was sent a reminder but unfortunately feedback was never received.

[7] The following studies included more than one treatment group and therefore allowed for multiple comparisons between treatment and control groups: Barrett (1998), Barrett et al (1996), Flannery-Schroeder & Kendall (2000), Mendlowitz et al (1999), Muris et al (1998), Öst et al (2001), Silverman, Kurtines, Ginsburg, Weems, Rabian, et al (1999), and Spence et al (2000). One comparison was excluded from the (Hayward et al., 2000) as the comparison group was a non-disordered sample.

foundation for moderator analyses. Statistical data from studies were entered separately into meta-analysis software for analysis (Borenstein & Rothstein, 2001).

2.5 Schmidt-Hunter Meta-analysis Methods

This meta-analysis followed the methods recommended by Schmidt-Hunter (1990) and Weiss and Weisz (1990). These methods were chosen as they have been developed in an attempt to address some of the major criticisms of meta-analyses. The key principles are: (a) a strong emphasis on effect size and making adjustments for sample size, instrument reliability and range restriction to increase accuracy;[8] (b) only one effect size from each study is included in the overall and moderator analyses; (c) effects that measure the same constructs or domains are combined; and (d) judgements are made about methodological quality. Additionally, the main research question was narrower than initial treatment outcome meta-analyses. In this meta-analysis, the focus was on treatment for just one group of disorders (see also Michael & Crowley, 2002; Rector & Beck, 2001).

2.6 Analyses

Statistical analyses were performed with the assistance of Comprehensive Meta-analysis™ (Borenstein & Rothstein, 2001) and Excel™. The software Comprehensive Meta-analysis™ allowed for the calculation of individual effect sizes, mean effect sizes, confidence intervals and *p*-values.

2.6.1 Missing Data

Seven authors were contacted to provide data that were missing from published papers. These data included sample sizes, means and standard deviations for outcome measures, and descriptive information for coding. Four (57%) authors replied and provided the missing data. This left two studies with missing descriptive data (Muris et al., 1998; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999) and one with missing data required for the calculation of

[8] In this meta-analysis, effect sizes were not adjusted for instrument reliability, differences in reliability across studies and other artifacts Schmidt-Hunter methods suggest may influence the effect size. As Hunter and Schmidt (1990) note the information required for these adjustments is seldom available in published studies. The method of Weiss and Weisz (1990) was followed here, where they felt making estimations would be unsystematic and potentially cause more bias than it would avoid. Therefore Schmidt-Hunter reliability adjustments were not made.

effect sizes (Silverman, Kurtines, Ginsburg, Weems, Lumpkin, et al., 1999).

As the majority of authors provided missing data, estimations for effect sizes using reported F, t and χ^2 statistics was not necessary. In cases where nonsignificant effects were reported without any statistic or data could not be converted to an effect size, the effect was estimated at 0.00.[9] Although conservative, this method is commonly used including in benchmark meta-analyses (e.g., Durlak et al., 1991; Michael & Crowley, 2002; Weisz et al., 1995). This was done in the one study (7 within-group follow-up effect sizes) where data were not available from the author (Silverman, Kurtines, Ginsburg, Weems, Lumkin, et al., 1999).

2.6.2 Calculation of Effect Sizes

As stated earlier, an aim of this meta-analysis was to assess and compare two methods of calculating effect sizes. The first method was the convention where posttreatment scores are used as the unit of measurement. The second method was to calculate effect sizes based on gain scores (change in score between pretreatment and posttreatment). Both methods involve comparisons between treatment and control groups.

Typically, authors include posttreatment means and standard deviations in published articles thus making data available for the calculation of effect sizes based on posttreatment scores. However, reporting of gain scores is much less common, and of those included studies in this meta-analysis, none reported gain score means and standard deviations. Authors were contacted and asked to provide gain scores for treatment and comparison groups (see also Clarke & Oxman, 2002). This decision was based on recommendations in the meta-analysis literature, where it is preferable to obtain raw data (in this case, gain scores) from authors, as opposed to estimating data, to maximise the accuracy of calculations (see Clarke & Oxman, 2002; Hunter & Schmidt, 1990).

Thirteen authors were contacted via email.[10] The nature of the research project was explained to them and gain scores were requested. See Appendix 1 for a copy of this email. In four cases, second authors were contacted in an effort to

---

[9] This method is conservative, particularly if the estimated effect size is in the direction of other effect sizes (i.e. positive) (Meyer & Marks, 1995).

[10] This differs from the N of 17 studies as Barrett, Kendall, and Silverman were contact authors for more than one study. As the Ginsburg & Drake (2002) study was published and included later in the year after the decision had been made to omit gain scores from effect size calculations, it was not necessary to request data for this study.

locate the corresponding author. This was successful in three cases (Cornwall et al., 1996; Mendlowitz et al., 1999; Shortt et al., 2001). The corresponding author for one study (Last et al., 1998) was not able to be contacted despite contact being made with previous employers and colleagues.

Seven authors replied to the original email request. Two (Muris et al., 1998; Spence et al., 2000) were unable to supply the data as it was inaccessible due to software difficulties, and two responded they were too busy to provide the data (Cornwall et al., 1996; Kendall, 1994; Kendall et al., 1997). Three agreed to provide the data (Beidel et al., 2000; Flannery-Schroeder & Kendall, 2000; Shortt et al., 2001). Authors that had agreed to provide data (n = 3) and those who had not responded (n = 5) were sent reminders approximately 2 & 4 months later. This resulted in data from five authors (Barrett, 1998; Barrett et al., 1996; Beidel et al., 2000; Hayward et al., 2000; Mendlowitz et al., 1999; Öst et al., 2001). Three authors also agreed to provide the data at a later date but unfortunately data were never made available (Flannery-Schroeder & Kendall, 2000; Shortt et al., 2001; Silverman, Kurtines, Ginsburg, Weems, Lumpkin, et al., 1999; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999). The result was gain score data for 5 of the 17 included studies (29%),[11] and approximately 40 effect sizes (including gain scores based on follow-up points). The low response and the resulting small number of gain effect sizes on which to base analyses and comparisons led to the method for this meta-analysis to be reconsidered. It was also of concern that a number of the key researchers (e.g. Kendall and Silverman) had not been able to provide data.

The decision was made to relinquish the calculation of between-group gain effect sizes. Unfortunately, this meant one of the initial aims of the study could not be met. Instead outcomes would be measured by within-group gain effect sizes and between-group posttreatment effect sizes. This method is consistent with a number of recent meta-analysis assessing treatment outcome (e.g. Michael & Crowley, 2002; Rector & Beck, 2001; Westen & Morrison, 2001).

The inclusion of within-group gain score effect sizes, possible from the published data, would provide an indication of the magnitude of change for treatment and control conditions. The effect sizes resulting from the two methods of calculation were not combined as some (e.g., Lipsey & Wilson, 2001) have

[11] Data for two of the studies had been sent through as raw data, necessitating the calculation of gain scores.

suggested in the meta-analysis literature. Within-group effect sizes have been shown to be an inflated estimate of effect due to the lack of comparison with a control (Lipsey & Wilson; Michael & Crowley, 2002. Thus, the two types of effect sizes are reported separately, following the procedure of recently published meta-analyses using this combined methodology (e.g. Michael & Crowley; Rector & Beck, 2001; Westen & Morrison, 2001).

In line with these other studies between-group effect sizes were considered to be the main effects for this meta-analysis and within-group effect sizes were reported as a useful addition.

2.6.3 Between-group Effect Sizes

Treatment and control group comparison were expressed using Hedges adjusted *g* (Hedges & Olkin, 1985). This effect size index was chosen as the standardised effect size (*d*) has been shown to be upwardly biased when studies have small sample sizes, particularly those less than 20 (Lipsey & Wilson, 2001). This meta-analysis only included one study with a sample size less than 20 (Ginsburg & Drake, 2002). However, as the decision was made to adjust the sample size of the control group based on the number of comparisons in the study, this decreased the sample size for a number of studies. This was done in an attempt to address the issue of the non-independence of effect sizes due to multiple comparisons within studies (Meyer & Mark, 1995). The use of *g* also followed the methods of Weisz and colleagues (1995).

The pooled standard deviation was used as the denominator as it is considered the most precise estimate of population variance and thought to have less sampling error (Hunter & Schmidt, 1990).[12] The following formulas were used to calculate Hedges adjusted *g*, the standard error and weight, in Comprehensive Meta-analysis™:

$$g = \left(\frac{m_1 - m_2}{s_p}\right)\left(1 - \frac{3}{4N - 9}\right)$$

[12] We chose to follow the method of Hunter and Schmidt (1990) who suggest there is rarely a difference between the standard deviation of the control and treatment group. Thus it is reasonable to use the statistic with the least amount of sampling variance (i.e. pooled standard deviation).

$$s_p = \sqrt{\frac{(n_1 - 1)SD^2_1 + (n_2 - 1)SD^2_2)}{n_1 + n_2 - 2}}$$

$$SE = \sqrt{\frac{N}{(n_1\, n_2)} + \frac{g^2}{2(N - 3.94)}}$$

$$w = \frac{1}{SE(\Theta)^2}$$

Where: m_1 = mean of the treated group
m_2 = mean of the control group
s_p = pooled within-group standard deviation
N = total sample size ($n_1 + n_2$)

For the 17 studies, between-group effect sizes were calculated comparing active treatment versus a no treatment (n = 1), waitlist (n = 11), or a placebo control (n = 5).[13] There were a total of 246 between-group effect sizes calculated at posttreatment. Twenty-one between-group effect sizes were also calculated for the studies (n = 3) reporting relevant follow-up data.[14] This allowed for a calculation of a treatment maintenance metric and was a further measure of effectiveness. If a study reported more than one follow-up assessment, only the follow-up with the longest duration was used. This was done to avoid weighting toward studies with numerous follow-up points (Kazdin et al., 1990).

Positive effect sizes indicated an effect in the expected direction (i.e., the treatment group did better than the control group) and a negative effect size indicated the treatment group fared worse than the control group (Lipsey & Wilson, 2001; Rosenthal, 1994). Eight percent of the between-group effect sizes calculated at posttreatment were negative (n = 20), and 10% were negative at follow-up (n = 2).

Cohen (1988) guidelines were used in interpreting small (0.20), medium (0.50), and large (0.80) effect sizes.

[13] See Table 14.2 Kendall et al. (1999) for a useful description of types of control conditions employed in treatment outcome research.

[14] Cornwall et al. (1996), and Hayward et al. (2000), and Silverman Kurtines, Ginsburg, Weems, Rabian, et al. (1999) included follow-up assessments of control groups. This was not available for the other studies where controls were treated after posttreatment assessment.

2.6.4 *Within-group Effect Sizes*

Within-group treatment effects were calculated using standardised mean gain scores.[15] A standardised difference between time 1 and time 2 is used to calculate an effect size when different studies have used differing operationalisation of an outcome (i.e., different outcome measures). A pooled standard deviation was used in order to be consistent with between-group effect size calculations as recommended in Hunter & Schmidt (1990). The formulas are as follows (Becker, 1988; cited in Lipsey & Wilson, 2001):

$$ES_{gain} = \frac{m_{T1} - m_{T2}}{s_p}$$

$$s_p = \sqrt{(s^2_{T1} + s^2_{T2}) / 2}$$

$$SE_{gain} = \sqrt{\frac{2(1-r)}{n} + \frac{ES^2_{gain}}{2n}}$$

$$w_{gain} = \frac{1}{SE^2_{gain}} = \frac{2n}{4(1-r) + ES^2_{gain}}$$

$$r = \sqrt{\frac{d^2}{d^2 + 4}}$$

Where:

m_{T1}	=	group mean at time 1 (i.e., pretreatment)
m_{T2}	=	group mean at time 2 (i.e., posttreatment)
s_p	=	pooled within-group standard deviation
r	=	correlation between scores at time 1 and time 2

A total of 268 within-group effect sizes were calculated estimating gain between pretreatment and posttreatment (161 for treatment conditions and 107 for control conditions).

Within-group effect sizes were also calculated for pretreatment to follow-up intervals to assess the maintenance of treatment gains. This was useful as few between-group effect sizes could be calculated at follow-up due to the majority of the 11 studies with follow-up assessments later treating control group participants (n = 7, 64%). As with between-group effect sizes, if a study reported more than

[15] Calculations were made in Excel™ as it is not possible to calculate within-group effects in Comprehensive Meta-analysis™ (M. Borenstein, personal communication, 7 January 2003).

one follow-up assessment, as discussed, the follow-up with the longest duration was used. Within-group effect sizes were also calculated for control groups.[16]

For the studies reporting follow-up data, 122 effect sizes estimating gain between pretreatment and follow-up for treatment conditions were calculated. An additional seven effect sizes were calculated for the placebo control condition (Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999), and 13 effect sizes were also calculated for the two studies reporting long-term follow-up data (Barrett et al., 2001; Kendall & Southam-Gerow, 1996). Seven percent of the within-group effect sizes calculated at posttreatment were negative (n = 20), and less than 1% were negative at follow-up (n = 1).

2.6.5 Effect Sizes from Proportions

Effect sizes for outcomes reported as proportions or percentages were calculated using arcsine transformation procedures (Cohen, 1988).[17] The resulting effect size is defined as the difference between the arcsine of the percentage in the treatment group and that of the control group. Unlike the other effect size index, there is no adjustment for sample size (Cohen, 1988). Thus effect sizes based on percentages were reported separately.

This was done for 12 studies reporting the percentage of treatment and control participants diagnosis free at posttreatment and one at follow-up. Effect sizes were calculated with the formula for directional tests (one-tailed) and tables were used to transform percentages into arcsine values (see Cohen, 1988).

$$ES_{arcsin} = \Phi_1 - \Phi_2$$

$$\Phi = 2 \text{ arcsine } \sqrt{p}$$

Where: Φ_1 = arcsin of the treatment group
Φ_2 = arcsin of the control group
p = observed proportion

[16] This was also done by Michael and Crowley (2002). Additionally, we sought to increase knowledge of the outcomes of attention control condition which have been found to have reduced anxiety in children and adolescents (e.g. Last et al., 1998; Silverman Kurtines, Ginsburg, Weems, Rabian, et al., 1999).

[17] Lipsey and Wilson (2001) recommend this method as it is a conservative method of transforming percentages. This method was used in a recent treatment outcome meta-analysis (Prendergast, Podus, Chang, & Urada, 2002).

2.6.6 Multiple Outcome Measures

All studies had multiple outcome measures. To calculate the overall mean effect size, effect sizes were averaged within a study to give a single effect size per study so as not to weight studies according to the number of outcome measures.[18] This was done as it is important to maintain the statistical independence by including only one effect size per study in any analysis, and is consistent with both Schmidt-Hunter methods and benchmark studies (Casey & Berman, 1985; Durlak et al; 1990; Kazdin et al., 1990; Weisz et al., 1987; 1995).[19] If subscale and total scores were reported, effect sizes based on total scores were used.

The overall mean effect size across all studies included only those outcomes assessing anxiety. Meta-analysis literature suggests it is not appropriate to pool measures assessing different psychological constructs (Wolf, 1986). A number of benchmark studies (e.g. Weiss & Weisz, 1990; Wiesz et al., 1995) have not followed this procedure as they have included a heterogeneous group of disorders. In contrast, meta-analysis of treatment outcomes for a single disorder have calculated effect sizes from outcomes measuring the construct of interest (e.g. Michael & Crowley, 2002; Rector & Beck, 2001; Westen & Morrison, 2001).

In studies where there was more than one treatment condition a mean effect size was calculated for each of the individual treatments. If studies presented more than one effect size for a construct or domain of functioning, an overall mean effect size was calculated. Thus, despite some variation in how individual studies reported findings, each study only contributed one effect size to the overall effect size calculation (Lipsey & Wilson, 2001).

---

[18] There are other methods of choosing an effect size, such as randomly selecting one effect size per study or treatment condition, selecting the median effect size, or giving priority to the most widely used measures (Davidson & Parker, 2001; Weiss & Weisz, 1990; Westen & Morrison, 2001). It has been noted that the means of selection can be somewhat unsystematic and there is evidence that results are dependent on the particular effect selected (see Weiss & Weisz, 1990). When these methods of calculating an overall mean effect size were compared, comparison showed no significant difference. Therefore, in keeping with Hunter and Schmidt (1990) and benchmark studies we chose to use the mean effect size from each study as this is a more systematic means of selection.

[19] Hunter and Schmidt (1990) add that including multiple effects from the same sample is problematic, as it is not consistent with assumptions that each effect size is based on an independent sample. However, Weiss and Weisz (1990) found a significant relationship between the number of effect sizes and within-study effect sizes. Studies with more effect sizes produced smaller effects.

2.6.7 *Outliers*

The impact of outliers was assessed using the definition from Weisz and colleagues (1995) (i.e., within one *SD* from an adjacent effect size). To identify outliers and examine the relationship between magnitude of effect size and sample size, effects were plotted using a funnel plot (Hunter & Schmidt, 1990; Borenstein & Rothstein, 2001). The standard deviation of effect sizes was 0.51 and no outliers were identified.

2.6.8 *Evaluation of Clinical Significance*

Two methods were used to examine the clinical significance of treatments. A requirement of these methods is a common outcome measure across treatments and studies. As recommended by Sheldrick and colleagues (2001), the outcome measure ought to be widely used and have sound reliability and normative data. Of the 17 studies, the most commonly used measure was the Revised Children's Manifest Anxiety Scale (RCMAS), with 11 studies (65%) including this self-report measure. The Child Behaviour Checklist (CBCL) was used in nine studies (53%).[20]

The RCMAS and CBCL were also chosen because they have normative data and report high levels of internal reliability (Reynolds & Richmond, 1985; cited in Boyd et al., 2000; Achenbach, 1991; cited in Sheldrick et al., 2001). Both of these measures were used by a number of studies in the current sample to assess clinical significance (e.g. see Barrett et al., 2001; Flannery-Schroeder & Kendall, 2000; Kendall, 1994; 1997; Shortt et al., 2001; Silverman, Kurtines, Ginsburg, Weems, Lumpkin, et al., 1999; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al, 1999). When studies included parent report versions of the RCMAS as well as child reports, the child report was used.

Thirteen studies were identified that either used the CBCL-I or RCMAS.[21] The following treatments were assessed, individual CBT, CBT plus family

[20] Studies reported scores for both the Internalising (CBCL-I) and Externalising (CBCL-E) of the CBCL, but the CBCL-I only was used as it contains the Anxious/Depressed scale, which is of most relevance to this meta-analysis.

[21] Included studies were: Barrett (1998); Barrett et al (1996); Cornwall et al (1996); Flannery-Schroeder & Kendall (2000); Kendall (1994); Kendall et al (1997); Mendlowitz et al (1999); Öst et al (2001); Silverman Kurtines, Ginsburg, Weems, Lumpkin, et al (1999); Silverman Kurtines, Ginsburg, Weems, Rabian, et al., (1999); and Spence et al (2000). Unfortunately the following

involvement, group CBT, family group CBT, parent only CBT and emotive therapy. Consistent with Sheldrick and colleagues (2001), if a study had CBCL data from mothers and fathers, only data from mothers were used.

2.6.8.1 Reliable Change Index (RCI)

This method was adapted by Abramowitz (1998) for group data and is used to evaluate clinically significant change. A RCI is calculated for each study on the outcome measure of interest using the following formulas (Jacobson & Truax, 1991):

$$RCI = \frac{M_{pre} - M_{post}}{S_{diff}} \qquad S_{diff} = \sqrt{2\,(SE)^2}$$

$$SE = SD_{pre}\ \sqrt{1 - r_{xx}}$$

Where:

M_{pre}	=	pretreatment score for the treatment group
M_{post}	=	posttreatment score for the treatment group
S_{diff}	=	standard error of difference between the two means
r_{xx}	=	retest reliability of the outcome measure

A RCI of 1.96 ($p < .05$) or greater shows the treatment change can be considered reliable and clinically significant (Jacobson & Truax, 1991).

2.6.8.2 Normative Comparison

This method was initially developed by Kendall & Grove (1988; cited in Kendall et al., 1999) to establish whether treated participants had returned to within a normative range on an outcome measure.[22]

The first step for this method in a meta-analysis is to conduct a two-tailed *z* test to establish whether the mean and the normative mean can be considered

---

four studies not included, as they did not use the CBCL-I or RCMAS: Ginsburg & Drake (2002); Hayward et al (2000); Last et al (1998); and Muris et al (1998).

[22] Normative range is defined as being within one standard deviation of the normative mean (Sheldrick et al., 2001). For the CBCl-I, the procedures of Kendall et al. (1999) were followed in establishing a normative comparison group. When analysing group data they recommend using the most conservative normative group. In this instance, they recommended boys aged 12 to 18 years as this norm group has the smallest sample size. For the RCMAS (Reynolds & Paget, 1983), the 12-year-old normative group was chosen.

clinically equivalent (equivalency testing). The following formula is (Kendall, Marrs-Garcia, et al., 1999):

$$z = \frac{M_N - M_C - \delta}{\sqrt{\left(\frac{(n_N - 1)SD^2_N + (n_C - 1)SD^2_C}{n_N + n_C - 2}\right)\left(\frac{1}{n_N} + \frac{1}{n_C}\right)}}$$

Where:

M_N	=	mean of the normative group
M_C	=	mean of the treatment group
δ	=	specified one standard deviation range of closeness
SD_N	=	standard deviation of the normative group
SD_C	=	standard deviation of the treatment group
n_N	=	sample size of the normative group
n_C	=	sample size of the treatment group

A z of 1.64 ($p < .05$) or greater shows the treatment mean is likely to be within the range of closeness to the normative mean (Sheldrick et al., 2001) and suggests the treated group is not meaningfully different from the normative group.

The next step is to conduct a traditional t test to establish whether differences exist between group means.[23] The results of the two tests are then combined and classified in 4 cells as statistically different/clinically equivalent, clinically equivalent, different but not clinically equivalent, or equivocal findings with more power required (Kendall, Marrs-Garcia, et al., 1999). Results that are classified as clinically equivalent or statistically different and clinically equivalent indicate clinically significant findings.

2.6.9 Homogeneity

Homogeneity analysis (Q) was conducted to assess whether effects differed more than would be expected due to sampling error (Hunter & Schmidt, 1990). The hypothesis that the studies are all drawn from the same population (i.e. from a population with the same effect size) is rejected if Q is found to have an significant p-value.

Borenstein and Rothstein (2001) recommend that a non-significant p-value is not in itself evidence of homogeneity, as this may reflect low power. Q has low

[23] This differs from the equivalency z test which established whether the difference between the treated groups mean and the normative group were within a given range.

statistical power for rejecting homogeneity when there are few effect sizes and studies in a meta-analysis (Lipsey & Wilson, 2001). Hedges and Olkin (1985) also warn that it is rare to have a truly homogeneous group of effect sizes.

As this meta-analysis had a small sample (N = 17), and there was low power to safely reject heterogeneity, the random effects model was adopted. The random model is more conservative as it tends to have smaller effect sizes and wider confidence intervals. This is because it reflects both variation within and between studies (Lipsey & Wilson, 2001). The random effects model assumes that samples are drawn from populations with different effect sizes and that effect sizes differ due to characteristics of the studies. Consequently, the advantage of the random effects model is increased external validity.

Q was calculated for the overall mean effect size and effect sizes in the moderator analyses (Hedges & Olkin, 1985) using the following formula:

$$Q = \sum w_i \ (\Theta - \Theta)^2$$

$$w_i = \frac{1}{SE(\Theta)^2}$$

2.7 Moderator Analyses

If effect sizes were found to be heterogeneous, moderator analyses were conducted to identify sources of variation in effect sizes. As recommended by Hunter and Schmidt (1990), variables were assessed to see if they were moderators by categorising them and comparing the associated mean effect sizes. A moderator is observed in a large difference between the mean effect sizes of two categories, a reduction in variance within the categories (Q), and where confidence intervals do not overlap (Borenstein & Rothstein, 2001; Hunter & Schmidt, 1990).

Moderators of interest included:

Sample characteristics: Consistent with Weisz and colleagues (1995), studies were categorised into those where most participants were male or female and on the basis of whether they included adolescent participants (13-years-old and over) or included child participants only (ranging in age up to 12-years-old). Whether participants were diagnosed with a specific anxiety disorder or participants with a

range of disorders were also categorised, to address the question of whether homogenous groups are associated with better outcomes.

Outcome measures: Previous meta-analyses (e.g. Weisz et al., 1995) have suggested assessment modalities are associated with differential effect sizes. Different modalities were grouped for moderator analyses (self-report, parent report, teacher report, independent observation, blind clinician ratings, and therapist ratings). The constructs assessed were anxiety and fear, cognition, social skills, depression symptoms, and personality.

Treatment approach: Studies were categorised based on the mode of treatment (i.e. group versus individual) and whether family and parents were included in treatment.

Methodology: As discussed previously, various authors have suggested the quality of studies can produce differences in effect sizes and thus affect the findings of a meta-analysis (e.g. Hunter & Schmidt, 1990). To examine whether methodological factors impacted on treatment outcome, studies were coded based on design features following the method of Durlak et al (1991). Studies received a point for each criterion met, with scores ranging from 0 to 6. Criteria were: (a) sample size $\geq$ 30; (b) random assignment; (c) low attrition (less than 10% and equal across groups); (d) included at least one normed or blinded behavioural outcome measure; (e) included an attention placebo control group; and (f) reported all posttreatment data for all instruments used in the study (including re-administering of diagnostic measures post-treatment). The type of control group was also coded to assess whether studies with attention control groups produced different effect sizes as compared to no treatment or waitlist controls.

2.8 File Drawer Analysis

File drawer analysis attempts to address type I publication bias, the error of finding a more positive result than really exists. This error results from what has been termed by Rosenthal (1979), the 'file-drawer problem', where the tendency is for studies producing non-significant results to remain unpublished and therefore unidentified for inclusion in a meta-analysis. The formula, developed by Hunter and Schmidt (1990) and Orwin (1983; cited in Lipsey & Wilson, 2001), was used to calculate the number of studies needed (supporting the null hypothesis) to reduce an effect size to a specified level. The formula is as follows (Orwin):

$$k_0 = k \left[\frac{ES_k}{ES_c} - 1 \right]$$

Where: k = the total number of studies in the mean effect size
ES_k = the weighted mean effect size
ES_c = the criterion effect size level (i.e., 0.20, 0.50, 0.80)

2.9 Power Analysis

Power analysis was conducted to establish the probability of failing to detect a treatment effect (i.e. type II error, the probability of failing to reject a false null hypothesis). Power values were calculated using a significance criterion set at $\alpha = .05$, sample size and Cohen's (1988) directional (one-tailed) power tables. The test was one-tailed as the hypothesis was $m_{treatment} > m_{control}$.

Conventional definitions of effect size were used to calculate power to detect small ($d = 0.20$), medium ($d = 0.50$), and large effects ($d = 0.80$). As control and treatment group sample sizes were not equal, the harmonic mean was calculated (Cohen, 1988):

$$n = \frac{2n_1 n_2}{n_1 + n_2}$$

3 Results

3.1 Descriptive Statistics

3.1.1 Characteristics of Participants

A summary of data for participants is presented in Table 1. The included studies (N = 17) produced a sample of 984 children and adolescents. The average study sample size was 57.9 (SD = 26.8). A total of 885 participants completed conditions: 571 participants were treated (m = 33.6, SD = 18.7), compared to 314 control participants (m = 18.5, SD = 8.5). Forty three percent of participants were male and 57% were female. Participants ranged in age from 6 to 17 years (m = 11.0, SD = 2.46). Thirteen studies included adolescents (13-17 years) in their sample. Eleven studies (65%) reported the ethnic make up of their sample. Seventy-one percent of the sample were Caucasian and 28% were from minority ethnic groups (16% African American,[24] 12% Hispanic, 0.4% Asian, and 2.9% other minority group). Only seven studies (41%) reported the socio-economic status of the participants. In these studies, 13% of the participants were considered to be from low-income households.

All studies reported primary diagnoses for participants. Twenty-four-percent of the children were diagnosed with social phobia; 19%, OAD; 19%, simple phobia; 16.4%, SAD; 12%, GAD; 3%, avoidant disorder; 1%, agoraphobia; 0.3%, panic disorder; and 0.2% OCD.[25] Percentages of participants with more than one anxiety disorder in the studies ranged from 22% to 89% (m = 45.9%, SD = 19.0).[26] The percentage of participants with comorbid disorders other than anxiety disorders ranged from 0% to 77% (m = 37.1, SD = 26.4). Simple phobia was most frequently cited as the most common comorbid anxiety disorder in eight studies; depression the most common other comorbid disorder in eleven studies. Only one study reported participants who were medicated in their sample (Mendlowitz et al., 1999). The authors stated that medications were maintained throughout treatment and therefore were not thought to confound treatment effects. All other studies excluded medicated participants (n = 14) or did not report this data (n = 2).

[24] It is noted the sample in Ginsburg & Drake (2002) were all were African-American youth.
[25] Percentages do not add to 100% due to differential reporting in studies.

Table 1: Summary of study participants

Characteristic	Frequency	*M*	*SD*	Range	*N* (studies)
Total number of participants	984	57.9	26.8	12-118	17
Total number of completers	885	52.1	25.1	9-102	17
% Male		43.2%		0-62%	16
Age		11.07	2.46	6-17	16
% From low socio-economic households		13.4%			6
Ethnicity					11
Caucasian	71.2%				
African American	16.4%				
Hispanic	12.4%				
Asian	0.4%				
Other	2.9%				
Total ethnic minority	28.4%				
Primary diagnosis					17
Social phobia	24.4%				
Simple phobia	19.3%				
OAD	19.0%				
SAD	16.4%				
GAD	12.4%				
Avoidant disorder	3%				
Agoraphobia	0.8%				
Panic disorder	0.3%				
OCD	0.2%				
Treated sample diagnosis free					
Posttreatment	63.7%			53.2-75.0%	12
Followup	73.5%			67.0-88.0%	9
Control sample diagnosis free					
Posttreatment	15.1%			4.0-56.0%	12
Followup	0.44%				1
> 1 anxiety disorder	81.3%				16
Other comorbid disorders	37.1%			0-77%	13
Medicated[a]	<.01%				15

Note: [a] Mendlowitz et al., (1999).

3.1.2 Characteristics of Studies

Summary data for the included studies are presented in Table 2 and the characteristics of the 17 included studies are presented in Table 3. The 17 studies included in this meta-analysis were published in peer reviewed journals between 1994 and 2002. Nine studies were conducted in the United States (53%), 6 in Australia (29%), 1 in Sweden (6%), 1 in Canada (6%), and 1 in the Netherlands

[26] Data was available for 14 studies as three studies excluded participants with more than one anxiety disorder and for one study information on the comorbidity of anxiety disorders could not be obtained from the author.

(6%). One study (Ginsburg & Drake, 2002) was conducted in a naturalistic school setting. All other studies (n = 16) and treatments were conducted in university departments or clinics associated with a university department.[27] Twelve studies (71%) recruited participants by advertising, 4 (24%) from clinic referrals, and one (6%) recruited participants at the school where the treatment was conducted.

Control participants (n = 314) were assigned to an attention placebo control condition in 5 studies (n = 70, 22%), a waitlist condition in 11 studies (n = 222, 71%), and a no treatment group in one study (n = 22, 7%). One control comparison group (Hayward et al., 2000) was excluded from the analyses as participants in the group did not have an anxiety disorder. Following posttreatment, control participants were treated in nine of 17 studies (53%).

Ninety-five percent of the treated participants (n = 541), were treated using CBT interventions; 2% (n = 12), emotive therapy; 1.6% (n = 9), behaviour therapy, and 1.6% (n = 9) EMDR therapy. Eight studies (47%) included treatment delivered in group mode, 7 studies (41%) included treatments with a family component, and one study (6%) included teachers in the treatment. All studies employing group or family treatments were CBT approaches. There were 26 treatment groups across all included studies. Seven studies compared two different treatments and one study compared three different treatments. Conditions included five different treatment modalities: 8 (31%) individual child therapies; 7 (37%), child group therapies; 8 (12%), individual family therapies; 2 (12%), family group therapies; and one (4%), a parent group therapy.

Four studies (24%), specifically targeted specific phobia; three (18%), social phobia; one (6%), school phobia; and the remaining nine targeted a heterogenous group of anxiety disorders (53%). All studies used a form of prescribed treatment or manuals. Eighty-eight percent (n = 15) of the studies reported the inclusion of homework in the treatment protocol. Homework exercises involved exposure or practice tasks for 9 of these studies (60%).[28] The mean duration for treatment was 12 sessions (range 1 – 24), and the mean length of sessions across the studies was 80 minutes (range of 40 to 180 minutes). Two studies also reported including booster sessions (Shortt et al., 2001; Spence et al.,

[27] Studies are considered efficacy studies as they were conducted in university settings, in contrast to effectiveness studies conducted in clinical and community settings (Lambert, 2001).

[28] Six studies did not report the nature of homework tasks and the remaining 2 studies were one day only treatments.

2000).

Fifteen studies reported data on therapists. In 9 (60%) of these studies, clinical psychologists were used as therapists; 5 (33%), employed doctoral candidates; 3 (20%), employed psychiatrists; 3 (20%), employed graduate students; 3 (20%) employed other professionals; and one (7%), employed a behavioural scientist. The mean number of therapists per study was 4.6 (*SD* = 2.86). The range was 1 to 11 therapists per study. The mean percentage of female therapists for the six studies reporting this data was 70%.

Eighty-eight percent of the studies included a measure of treatment integrity (n = 15). Forty-one percent assessed adherence to manuals (n = 7) and the remaining 47% used manual adherence and monitoring of treatment session via recording (i.e., video and audio) (n = 8). Nine studies (53%) reported using a treatment credibility measure at the end of treatment to rate client and parent satisfaction and perceived effectiveness of treatment. Fifty-three percent (n = 9) included a child credibility rating and 45% (n = 8) included child and parent ratings.

A multi-perspective assessment approach was adopted in all studies. In total, 51 different measures were used across the included studies (see Appendix 2). The number of measures per study ranged from one to 15 per study (m = 6.1, *SD* = 3.7). All studies included self-report measures, 88% (n = 15) included parent reports in the assessment, and 18% (n = 3) included teacher reports. Behavioural Avoidance Tests using independent observers were used by 7 studies (41%).

Diagnosis was established using a clinical interview in all the studies. Seventy-six percent of studies (n = 13), used versions of the Anxiety Disorders Interview Schedule - Children (ADIS-C); 12% (n = 2), Diagnostic Interview for Children and Adolescents-Revised (DICA-R); 6% (n = 1), Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS); and 6% (n = 1), the Diagnostic Interview Schedule for Children (DISC). Participants in 9 studies (53%) met DSM-III-R criteria and 8 (47%) DSM-IV criteria. Seventy-six percent (n = 13) used a blind rater to establish diagnosis. Thirteen studies (76%) used diagnostic status met as an outcome indicator. One study reported school attendance (Last et al., 1998) as an outcome and another study (Beidel et al., 2000) used self-reported number of negative coping and stressful events. Twelve studies used depression measures, most commonly the CDI (53%, n = 9), and one

study (6%) included a personality inventory (Beidel et al., 2000).

Fourteen studies (82%) included follow-up assessments. Follow-up periods ranged from 4 weeks to 6 years. The most common follow-up period was 12 months (n = 10). One study (7%) included follow-up assessment at 4 weeks, 4 studies at 3 months (29%), 5 studies at 6 months (36%), 10 studies at 12 months (71%), and 2 studies included long-term follow-up from 1 to 6 years (14%). Five studies (36%) included more than one follow-up interval.

Scores for methodological quality indicated the studies were well controlled, with a mean of 4.2 (SD = 0.56) on the 0 to 6 scale (Durlak et al., 1991). All studies randomly assigned participants and used at least one normed or blind outcome measure. Five studies (29%) met criteria for an attention placebo control; 5 (29%), attrition less than 10%; 14 (82%), reported all post-test data; and 14 (82%), a sample size greater than 30.

3.2 Main Effects

A total of 267 between-group effect sizes were calculated. All effect sizes were calculated based on control versus treatment group comparisons. Ninety-two percent of the effect sizes were calculated from post-treatment data (n = 246) and 8% from follow-up data (n = 21). The 246 post-treatment effect sizes are the basis of the main analysis presented first. Within-group effect sizes (n = 410) will be discussed separately.

The overall weighted mean effect size (g) comparing control and treatment groups post-treatment was 0.77 (95% CI 0.55-0.99, p <.001). The confidence interval indicates a 95% probability the overall mean effect size the falls between 0.55 and 0.99 and as the interval does not contain 0, the effect size is statistically significant (α = .05) (Lipsey & Wilson, 2001).

The overall effect size of 0.77 suggests the average treated participant was 0.77 standard deviation units above the average control group participant. The effect size is associated with a U_3 value of 78.8 (Cohen, 1988). This figure provides another indication of the size of the difference between the treatment and control group means. Seventy-nine percent of the treated group were more improved compared to the control group.

Mean effect sizes for each study ranged from 0.21 to 2.50 (α = .05). See Figure 1 and Table 4 for mean effect sizes, confidence intervals and p-values for

the included studies. No study produced an overall negative mean effect size and nine studies produced statistically significant results at $p < .05$. The mean effect size for CBT treatments was similar to the overall effect size ($g = 0.76$; $n = 15$; 95% *CI* 0.53-1.00; $p < .001$). Between group effect sizes calculated at follow-up produced an overall g of 0.65 ($n = 3$; 95% *CI* 0.42-0.87) for studies reporting data for treatment and control groups at follow-up assessment.

Analysis of homogeneity resulted in a Q-value of 33.12 with 16 degrees of freedom ($p < .01$). The significant p-value indicates the Q-value is greater than the critical chi-square value ($\alpha = .05$). This suggests heterogeneity of effect sizes, thus there was more variation amongst effect sizes than would be expected from sampling error alone. Examination of the dispersion of effects sizes on the Forrest plot confirmed variation of effect sizes, and also indicated all studies had beneficial treatment effects. Heterogeneity of effect sizes suggests a random effects model is appropriate and the examination of moderators may account for variation.

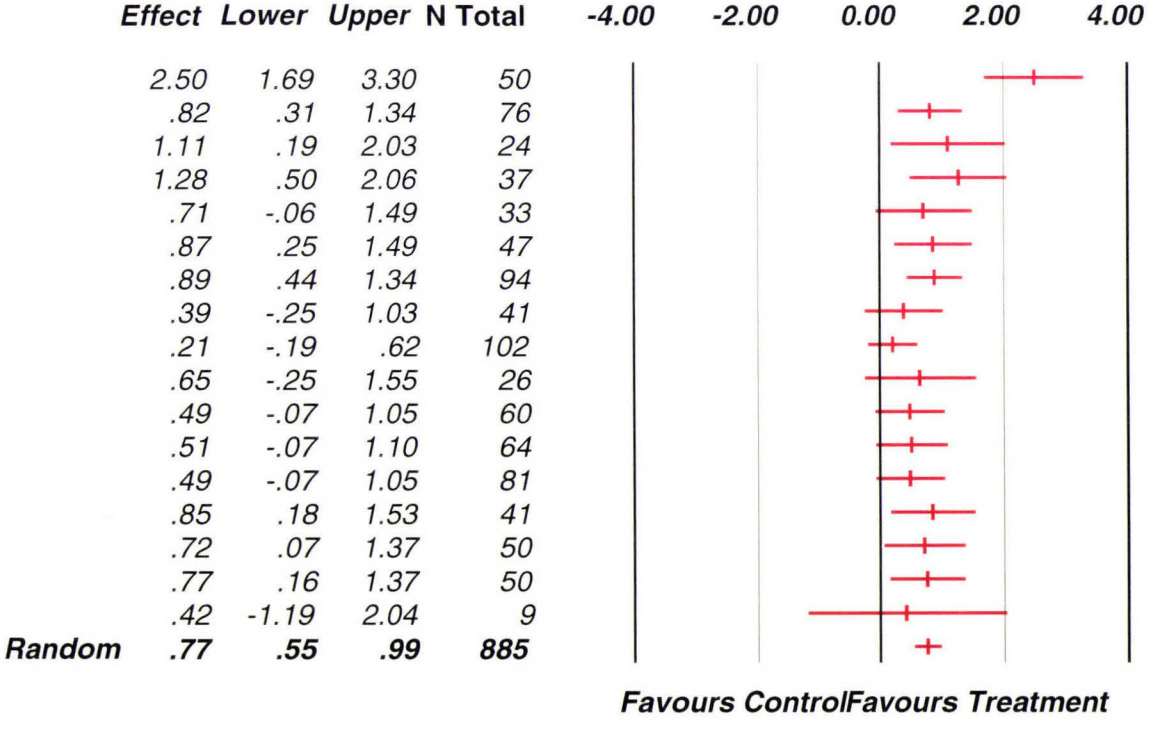

Figure 1: Forrest plot of effect sizes (g) for all studies

Table 2: Summary of included studies

Characteristic	Frequency	Percentage[a]	*M*	*SD*	Range
Year of publication					
1990-1994	1	5.9			
1995-1999	9	52.9			
2000-2002	7	41.2			
Treatment approach					
CBT	15	88.2			
Behaviour therapy	1	5.9			
EMDR	1	5.9			
Emotive therapy	1	5.9			
Treatment conditions	26				1-3
Individual (child)	8	47.1			
Group (child)	7	41.2			
Family	8	47.1			
Family group	2	11.8			
Parent group	1	5.9			
Family included in treatment	7	41.2			
Teacher included in treatment	1	5.9			
Treatment sessions					
Number			11.8	5.9	1-24
Length (minutes)			81.2	34.0	40-180
Treatment included homework	15	88.2			
Booster sessions	2	11.8			
Prescribed or manualised treatment	17	100			
Measure of treatment credibility	9	52.9			
Measure of treatment integrity	15	88.2			
Treatment manual adherence	7	41.2			
Treatment manual adherence and recorded monitoring	8	47.1			
Control comparison					
Waitlist	12	70.5			
No treatment	1	5.9			
Attention placebo	5	29.4			
Control group treated posttreatment	9	52.9			

Note: [a] Percentages do not add to 100%, as some studies did not report all data.

Characteristic	Frequency	Percentage	*M*	*SD*	Range
Assessment measures			6.1	3.7	2 – 15
Child self report	17	100			
Parental report	15	88.2			
Teacher report	3	17.6			
Behavioural Avoidance Test	5	29.4			
Blind assessment posttreatment	13	76.5			
Depression outcome	12	70.6			
Assessment of clinical significance	16	94.1			
Diagnosis free	13	76.5			
Normative comparison	7	41.2			
Clinician rated improvement	3	17.6			
Significant improvement from pretreatment	2	11.8			
Follow-up assessment	14	82.4			
None	3	17.6			
4 weeks	1	7.1			
3 months	4	26.8			
6 months	5	35.7			
12 months	10	71.4			
> 12 months	2	14.3			
Sample at follow-up (longest)	13	77.5			
Therapists			4.6		1-11
Clinical psychologists	9	60.0			
Psychiatrists	3	20.0			
Doctoral candidates	5	33.3			
Graduate students	3	20.0			
Behavioural scientist	1	6.7			
Other professionals	3	20.0			
Methodological quality (score 0-6)		(Mode = 4)	4.2	0.56	2-5
Sample size ≥ 30	14	82.4			
Random assignment	17	100			
Attrition ≤ 10%	5	29.4			
Normed or blinded outcome	17	100			
Attention placebo control	5	29.4			
Reported all post-test data used pre-test	14	82.4			
Attrition		13.3%		0-27%	17

Table 3: Characteristics of included studies

Study	Journal	Population Primary Diagnosis	Treatment condition/s	Control condition	*N* Control	*N* Treatment	Attrition
Barrett (1998)	JCCP	OAD, SAD, SOP	CBT Group & CBT Family Group	Wait list	16	39	17%
Barrett et al (1996)	JCCP	OAD, SAD, SOP	CBT Individual & CBT Family	Wait list	23	53	11%
Barrett et al (2001) [a]	JCCP	OAD, SAD, SOP	CBT Individual & CBT Family	Wait list	-	52	33%
Beidel et al (2000)	JAACAP	SOP	CBT Group	Attention control	20	30	17%
Cornwall et al (1996)	Behaviour Change	SP	Emotive Therapy	Wait list	12	12	0
Flannery-Schroder & Kendall (2000)	JAACAP	GAD, SAD, SOP	CBT Individual & CBT Group	Wait list	12	29	12%
Ginsburg & Drake (2002)	JAACAP	GAD, SP, SOP	CBT Group	Education Support	4	5	25%
Hayward et al (2000)	JCCP	SOP	CBT Group	No treatment	22	11	6%
Kendall (1994)	JCCP	OAD, SAD, AD	CBT Individual	Wait list	20	27	22%
Kendall et al (1997)	JAACAP	OAD, GAD, AD, SP, SAD	CBT Individual	Wait list	34	60	20%
Kendall & Southam-Gerow (1996) [b]	JCCP	OAD, SAD, AD	CBT Individual	Wait list	-	36	18%
Last et al (1998)	JAACAP	School Phobia	CBT Individual	Education Support	21	20	27%
Mendlowitz et al (1999)	JCCP	Anxiety disorders	CBT Group, CBT Family Group & CBT Parent Group	Wait list	40	62	9%
Muris et al (1998)	JCCP	SP	BT & EMDR	Attention control	8	18	-
Ost et al (2001)	Journal of Clinical Child Psychology	SP	CBT Individual & CBT Family	Wait list	19	60	0
Shortt et al (2001)	JCCP	SAD; GAD; SOP	CBT Family (FRIENDS)	Wait list	16	63	10%
Silverman et al (1999)	JCCP	SP, SOP, AG	CBT Family (Self Control & Contingency Management)	Education Support	16	65	22%
Silverman et al (1999)	Journal of Clinical Child Psychology	GAD, SOP, OAD	CBT Group	Wait list	16	25	27%
Spence et al (2000)	JCCP	SOP only	CBT Group & CBT Family Group	Wait list	14	36	-

Note: JCCP = Journal of Consulting & Clinical Psychology; JAACAP = Journal of the American Academy of Child & Adolescent Psychiatry; AG = Agoraphobia; AD = Avoidant Disorder; GAD = Generalised Anxiety Disorder; OAD = Overanxious Disorder; SAD = Separation Anxiety Disorder; SP = Simple/Specific Phobia; SOP = Social Phobia; CBT = Cognitive-Behavioural Therapy; EMDR = Eye Movement Desensitisation & Reprocessing; BT = Behavioural Therapy; [a]Follow-up study of Barret et al (1996); [b]Follow-up study of Kendall (1994).

Table 4: Summary of effect sizes across studies

Author	Treatment Modality	N		Effect	95% Confidence Intervals		Within-group	Confidence Intervals		Effect size from diagnosis free %
		Treatment	Control	(g)[a]	Lower	Upper	Effect[b]	Lower	Upper	
Barrett (1998)	CBT Group format	34	16	**2.50*** **	1.69	3.30	**2.47**	2.08	2.85	0.86
Barrett et al (1996)	CBT Family format	53	23	**0.82*** **	0.31	1.34	**0.82**	0.67	0.98	0.97
Barrett et al (2001)[c]	CBT Family format	52	-	-	-	-	**1.24**	1.05	1.43	-
Beidel et al (2000)	CBT Group format	30	20	**0.77****	0.16	1.37	**0.90**	0.74	1.05	>.995
Cornwall et al (1996)	Emotive Imagery	12	12	**1.11****	0.19	2.03	**1.19**	0.86	1.53	-
Flannery-Schroder & Kendall (2000)	CBT Group format	25	12	**1.28*** **	0.50	2.06	**1.05**	0.87	1.23	0.82
Ginsburg & Drake (2001)	CBT Group format	4	5	**0.42**	-1.19	2.04	**0.72**	-0.02	1.47	-
Hayward et al (2000)	CBT Group format	11	22	**0.71**	-0.06	1.49	**1.36**	0.90	1.83	0.86
Kendall (1994)	CBT Individual format	27	20	**0.87****	0.25	1.49	**1.12**	0.94	1.29	>.995
Kendall et al (1997)	CBT Individual format	60	34	**0.89*** **	0.44	1.34	**0.97**	0.87	1.07	>.995
Kendall & Southam-Gerow (1996)[d]	CBT Individual format	36	-	-	-	-	**1.13**	0.91	1.35	-
Last et al (1998)	CBT Individual format	20	21	**0.39**	-0.25	1.03	**0.61**	0.22	1.00	-
Mendlowitz et al (1999)	CBT-Family + Group format	62	40	**0.21**	-0.19	0.62	**0.18**	-0.06	0.42	-
Muris et al (1998)	EMDR & Behaviour Therapy	18	8	**0.65**	-0.25	1.55	**1.06**	0.75	1.37	-
Ost et al (2001)	CBT Individual format	41	19	**0.49**	-0.07	1.05	**0.48**	0.32	0.64	-
Shortt et al (2001)	CBT-Family + Group format	48	16	**0.51**	-0.07	1.10	**0.95**	0.74	1.16	>.995
Silverman et al (1999)	CBT+ Family component	65	16	**0.49**	-0.07	1.05	**0.74**	0.62	0.85	0.29
Silverman et al (1999)	CBT Group format	25	16	**0.85****	0.18	1.53	**1.02**	0.80	1.23	0.97
Spence et al (2000)	CBT Group format	36	14	**0.72***	0.07	1.37	**0.80**	0.63	0.98	0.98
	(17) Random effects	**571**	**314**	**0.77*** **	**0.55**	**0.99**	**0.88**	0.83	0.92	**>.995**

Note: [a] Calculated from the posttreatment means of treatment and control groups; [b] Calculated from the pretreatment and posttreatment means of treatment groups; [c] Follow-up study of Barret et al (1996); [d] Follow-up study of Kendall (1994); $*p < .05$; $**p < .01$; $***p < .001$.

3.2.1 Diagnostic Outcomes

For the 12 studies reporting diagnostic status of participants following treatment, the average number of the treated group that did not meet diagnostic criteria for their primary disorder post-treatment was 64% (range 45-75%), compared to 15% of controls (range 0.04 – 56%).[29] Eleven percent of waitlist and untreated controls ($n = 9$ studies) did not meet diagnostic criteria for their primary disorder following the control period, compared to 27% of the attention placebo controls ($n = 3$ studies). The mean effect size calculated from proportions (probit transformations) of treatment and control participants diagnosis free at posttreatment was 1.11 ($\alpha = .05$).

At follow-up, there was only one study with data for the control group (Hayward et al., 2000). The effect size calculated from proportions of treatment and control participants diagnosis free at follow-up was 1.45 ($\alpha = .05$). At the longest follow-up point (range 3-12 months), 74% of treated children did not meet diagnosis for their primary disorder (n = 9 studies). The two studies with long-term follow-up assessment (Barrett et al., 2001; Kendall & Southam-Gerow, 1996) reported 86% and 91% of treated participant's to be diagnosis free. Table 4 includes effect sizes calculated from proportions.

3.2.2 Within-group Effects

There were 410 within-group effect sizes calculated, with 268 effect sizes representing gain from pretreatment to posttreatment for treatment (161) and control (107) groups. The within-group mean treatment effect size across 17 studies was 0.88 (range 0.18-2.47; 95% *CI* 0.83-0.92). A mean within-group effect size was calculated for control groups to assess gain that could not be explained by treatment. The result was a mean effect size of 0.28 (range 0.00-0.92; 95% *CI* 0.21-0.35). Separate mean effect sizes were also calculated for attention placebo control groups versus waitlist control groups. The result was an attention mean effect size of 0.50 ($n = 5$; 95% *CI* 0.36-0.64), compared to a waitlist control mean effect size of 0.22 ($n = 12$; 95% *CI*

[29] The percentage of attention placebo control participants diagnosis free for Silverman Kurtines, Ginsburg, Weems, Rabian, et al. (1999) was considerably larger than the other studies (56%).

0.14-0.29).

To estimate within-group gain from pretreatment to follow-up 142 effect sizes were calculated. The resulting mean within-group effect size was 1.13 (range 0.92-4.04; 95% *CI* 1.08-1.19). The two studies which included long-term follow-up assessment produced a mean within-group effect size of 1.19 (95% *CI* 1.05-1.34). The one study including data at 12-month follow-up for an attention placebo control produced a within-group effect size of 1.21 (95% *CI* 1.10-1.33).

3.3 Clinical Significance

3.3.1 Reliable Change Index

The RCI method was used to evaluate whether change after treatment on two outcome measures was large enough to be considered clinically significant.[30] Of the 13 treatment conditions evaluated, five (38%) showed clinically significant change posttreatment as measured by the CBCL-I. See Table 5 for a summary of results for each treatment condition. Follow-up was associated with an increased number of conditions exhibiting clinically significant change. Seven of 12 conditions (55%) including follow-up assessment showed clinically significant change at follow-up. To summarise, all but one of those conditions showing clinically significant change at posttreatment maintained this at follow-up and a further two conditions showed significant change at follow-up when they had not at posttreatment. Four conditions did not show clinically significant change at posttreatment or follow-up.

3.3.2 Normative Comparisons

Equivalency testing was used to evaluate whether treatment conditions returned participants to within a normal range on the CBCL-I (range of closeness). Five conditions (38%) showed clinical equivalency on the CBCL-I, suggesting participants

[30] Normative comparisons using the RCMAS were removed from the analysis. This was done as initial comparison with the normative mean revealed treatment group means were not significantly different from the normative mean at pretreatment. Therefore it would be inaccurate to report results indicating treatment conditions had returned to within normal limits. This meant seven treatment conditions which had not used the CBCL-I were removed from analysis of clinical significance.

were returned to within normal limits at posttreatment. At follow-up, 12 conditions (83%) showed clinical equivalency. All conditions maintained clinical equivalency from posttreatment to follow-up. Only one condition did not show clinical equivalency at posttreatment or follow-up. This study also did not show significant change as measured by the RCI.

Table 5: CBCL-I, RCI and normative comparisons for each treatment condition

Study	Treatment Condition	RCI		Normative Comparison	
		Posttreatment	Follow-up	Posttreatment	Follow-up
Barrett et al (1998)	CBT Group format	*	*	Equivalent	Equivalent
Barrett et al (1996)	CBT Individual format		*	Different	Equivalent
Barrett et al (1996)	CBT Family format		*	Equivalent	Equivalent
Flannery-Schroder & Kendall (2000)	CBT Indivdual format	*	*	Equivalent	Equivalent
Flannery-Schroder & Kendall (2000)	CBT Group format	*		Different	Equivalent
Kendall (1994)	CBT Individual format	*	*	Different	Different
Kendall et al (1997)	CBT Individual format		*	Different	Equivalent
Shortt et al (2001)	CBT Family (FRIENDS)	*	*	Different	Equivalent
Silverman et al (1999)	CBT Family (SC)			Equivalent	Equivalent
Silverman et al (1999)	CBT Family (CM)			Equivalent	Equivalent
Silverman et al (1999)	CBT Group format		-	Different	-
Beidel et al (2000)	CBT Group format			Different	Different

Note: *Significant RCI (>1.96); Equivalent = significant clinical equivalency test and non-significant z-test; Different = non-significant clinical equivalency test and significant z-test.

3.4 Moderator Analysis

See Table 6 for a summary of moderator analyses using between-group effect sizes.

3.4.1 Interaction of Gender

Studies were analysed calculating a mean effect size using proportions of male

participants as the criterion (Weisz et al., 1995). The effect when the proportion of male participants was more than 50% was $g = 0.89$ ($n = 8$; 95% *CI* 0.69-1.10; $p < .01$). The effect when the proportion of male participants was less than 50% was $g = 0.49$ ($n = 8$; 95% *CI* 0.27-0.71; *ns*[31]). As there was a difference in the size of the mean effect sizes and as the confidence intervals only just overlap, a significant difference between the two groups is suggested. However *Q*-tests indicated heterogeneity and, given the small sample size, the result should be interpreted with caution (Borenstein & Rothstein, 2001; Hunter & Schmidt, 1990). The Forrest plot confirmed variation in the effects contained within the two groups of studies.

3.4.2 Interaction of Age

Studies were grouped by mean age using age 12 years as a cut off (Weisz et al., 1995). The result was 14 studies with a mean age of 12 years or younger and only one study with a mean age of 13 years or older.[32] It was not considered appropriate to conduct analysis with such a disproportionate numbers in the groups.

3.4.3 Interaction of Homogeneity of Sample

Studies were analysed by calculating a mean effect size for studies targeting one primary anxiety disorder and compared to studies with heterogeneous samples. The effect for studies with homogeneous samples was $g = 0.64$ ($n = 8$; 95% *CI* 0.41-0.87; *ns*) and for studies with heterogeneous samples was $g = 0.78$ ($n = 9$; 95% *CI* 0.59-0.97; $p < .001$). A further comparison of group treatment conditions also showed little difference between outcomes for groups targeting a particular diagnosis ($g = 0.73$; $n = 5$; 95% *CI* 0.41-1.06; *ns*) versus those groups with a range of diagnoses ($g = 0.71$; $n = 7$; 95% *CI* 0.14-1.29; $p < .05$). The small difference between the size of the mean effect sizes, the overlapping confidence intervals and the variation of effect sizes suggests homogeneity of the samples did not account for a significant difference in outcomes. The Forrest plot also confirmed considerable variation of effects within the groups.

---

[31] Non significant result.
[32] Mean age was missing for two studies.

3.4.4 Interaction of Assessment Modality

Mean effect sizes were calculated for data from self-report, parent report, observers and clinician ratings, and teacher reports of anxiety (Weisz et al., 1995). The results were as follows: self-report g = 0.50 (n = 137;[33] 95% *CI* 0.44-0.56; p <.001); parent report g = 0.73 (n = 75; 95% *CI* 0.65-0.81; p <.001); observer ratings g = 0.90 (n = 25; 95% *CI* 0.73-1.05; p <.001); clinician rating g = 1.71 (n = 6; 95% *CI* 1.38-2.04; p <.05); and teacher report g = 0.05 (n = 6; 95% *CI* -0.18-0.28; *ns*). There were considerable differences between mean effects sizes of these groups. Larger effects and smaller confidence intervals were associated with clinician ratings, observer ratings, and self-report and parent outcomes. By contrast, teacher reports were associated with very small effects. As the confidence intervals for self-report, parent report, observer ratings, and clinician ratings, also did not overlap and sample sizes were acceptable, results suggests these modes of assessment were associated with differential outcomes.

The mean effect sizes associated with different constructs assessed were also variable. The mean effect size for BATs was g = 1.19 (n = 8; 95% *CI* 0.95-1.42; p <.001); personality measures, g = 0.70 (n = 1; 95% *CI* 0.10-1.30; p = 1.00); depression measures, g = 0.57 (n = 10; 95% *CI* 0.39-0.74; *ns*); cognition measures, g = 0.46 (n = 4; 95% *CI* 0.02-0.90; p < .05); social skills measures, g = 0.33 (n = 3; 95% *CI* -0.03-0.70; *ns*); and coping measures, g = 0.13 (n = 4; 95% *CI* 0.008-0.54; p = .04). Again, there was considerable differences between mean effects sizes for the constructs assessed. The large confidence intervals reflect the variation amongst effects. Additionally, the sample sizes were small (n = 8), so it is not possible to conclude that the construct measured had a significant affect on effects. However, there is some evidence that BATs were associated with larger effects.

3.4.5 Interaction of Treatment Modality

Moderator analyses were conducted to look at the effect of family and group treatment modalities. When family were included in treatment conditions, the result was a mean effect size of g = 0.60 (n = 9; 95% *CI* 0.27-1.08; p = .001). The effect when family were

[33] *N* represents the number of outcomes.

not included was $g = 0.76$ ($n = 17$; 95% *CI* 0.57-0.96; *ns*). Although there was a small difference in the size of mean effect sizes, the confidence intervals do overlap and the Q values indicates variability of effects. This was confirmed on examination of the Forrest plot. It was thus not possible to conclude that the effect of including family had a significant effect on treatment outcomes.

Group format treatment conditions produced a mean effect size of $g = 0.62$ ($n = 12$; 95% *CI* 0.40-0.85; $p < .05$). The effect when treatment was conducted in individual format was $g = 0.77$ ($n = 14$; 95% *CI* 0.57-0.97; *ns*). The small difference in mean effect sizes, the overlap of confidence intervals and heterogeneity amongst the group format conditions mean it was not possible to establish that group format had a significant affect on treatment outcomes.

3.4.6 Interaction of Control Group

Studies which employed an attention placebo control group had a mean effect size of $g = 0.56$ ($n = 5$; 95% *CI* 0.25-0.87; *ns*). Those employing a waitlist control group produced a mean effect size of $g = 0.86$ ($n = 12$; 95% *CI* 0.57-1.14; $p < .001$). There was a difference in the size of the mean effect sizes and heterogeneity of effects was reduced when attention placebo studies were removed. While the variability amongst the attention placebo group and small sample size make it difficult to conclude whether there was an effect for type of control group, results do suggest differential effects with attention placebo producing smaller effects.

3.4.7 Effect moderated by Methodological Quality

Studies were separated into two groups based on the score they received during coding for methodological quality (0 – 6). As indicated in Durlak et al (1991) scores of greater than 5 indicate studies were well controlled and met almost all criteria. Therefore the cut offs were ≥5 and <5. Studies scoring ≥5 produced a mean effect size $g = 0.56$ ($n = 5$; 95% *CI* 0.29-0.83; *ns*) and studies with a score <5 produced a mean effect size of $g = 0.89$ ($n = 12$; 95% *CI* 0.63-1.20; $p < .01$). There was a difference in the size of the mean effect sizes. However the variability amongst the

studies, the overlap of confidence intervals and the small sample size meant it was not possible to establish if there was an effect for methodological quality.

Table 6: Mean effect size based on moderator analysis

Moderator variable	*N* (Studies)	Effect size *g*	95% Confidence Interval Lower	Upper	*Q*
Gender					
< 50% Male	8	0.49	0.27	0.71	6.63
> 50% Male	8	0.89	0.69	1.10	18.72*
Homogeneity of sample					
One primary diagnosis	8	0.64	0.41	0.87	2.91
Range of anxiety diagnoses	9	0.78	0.59	0.97	29.39***
Assessment Modality					
Self-report	137	0.50	0.44	0.56	248.50***
Parent report	75	0.73	0.65	0.81	197.53***
Observer rated	25	0.90	0.75	1.05	256.43***
Clinician rated	6	1.71	1.38	2.05	13.43*
Teacher report	6	0.05	-0.18	0.28	2.84
Construct assessed					
Behavioural Avoidance Test (BAT)	8	1.32	0.48	2.15	83.50***
Personality	1	0.70	0.10	1.30	.00
Depression	10	0.56	0.33	0.78	14.99
Cognition	4	0.46	0.02	0.90	9.57*
Coping	4	0.13	0.008	0.54	2.14*
Social skills	3	0.33	0.03	0.70	.40
Treatment modality					
Family included in treatment	9	0.60	0.35	0.85	19.35*
Family not included	17	0.76	0.58	0.94	18.29
Group format	12	0.62	0.40	0.84	21.00*
Individual format	14	0.77	0.58	0.97	16.71
Control comparison					
Waitlist/No treatment	12	0.77	0.60	0.94	30.81***
Attention placebo	5	0.56	0.25	0.87	.90
Methodological Quality					
< 5	12	0.81	0.63	0.98	30.16*
≥ 5	5	0.56	0.29	0.83	1.04

Note: $*p < .05$, $**p<.01$, $***p<.001$, for heterogeneity test (*Q*).

3.5 File Drawer Analysis

Using Orwin's formula (1983; cited in Lipsey & Wilson, 2001), 20 studies with effect sizes of zero would be need to reduce the treatment effect to a medium magnitude of 0.50. Forty-eight studies would be needed to reduce the treatment effect to a small magnitude of 0.20.

3.6 Power Analysis

Post-hoc analysis of power with the sample size of 885 (treatment n = 571; control n = 314) and using a directional test at the α = .05 level showed power exceeded the 80% criterion for detecting small, medium and large effects. Power was >.995 for the effect sizes of medium and large proportions and 0.81 for effects of small proportions. As there was a greater than 99% chance of detecting medium and large effects and an 81% chance of detecting small effects, it is unlikely that a type II error was made.

4 Discussion

4.1 Overall Treatment Effects

Behavioural and cognitive-behavioural interventions have been identified as efficacious treatments for children and adolescents with anxiety disorders (Chambless & Ollendick, 2001). The results of this meta-analysis confirm the efficacy of these treatments in an aggregated form. Furthermore, the effects from these studies suggest behavioural and cognitive-behavioural interventions are associated with positive outcomes for children and adolescents diagnosed with a range of different anxiety diagnoses and across a wide range of age groups.

Findings are comparable to past meta-analyses on the effectiveness of psychotherapy with children and adolescents. The overall between-group mean effect size for this review (0.77) is consistent with effects from previous meta-analyses ranging from 0.71 to 0.88 (Casey & Berman, 1985; Durlak et al., 1991; Kazdin et al., 1990; Michael & Crowley, 2002; Weiss & Weisz, 1990; Weisz et al., 1987; 1995). The overall effect size of 0.77 suggests the average treated child was better off than approximately 79% of untreated children. Cohen's (1988) guidelines for interpreting the magnitude of effect sizes places the overall mean effect size at the upper limit of the moderate range (between 0.50 and 0.79). The mean within-group gain effect size (0.88) across studies provides further support for the effectiveness of treatments with an effect size magnitude in the large range (greater than 0.80).

Overall, treated participants also had a greater chance of being diagnosis free. The mean effect size (>1.00) calculated from diagnostic status indicated treated children were much less likely than controls to meet criteria for their primary disorder post-treatment. On average, 64% of the treated group did not meet criteria for their primary diagnosis after treatment, compared to 15% of those in the control group.

Moreover follow-up data suggested treatment effects were maintained over time (range 3 to 12 months). The mean between-group effect at follow-up was 0.65. Unfortunately, this effect size was based on only three studies (Cornwall et al., 1997; Hayward et al., 2000; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999). However, the mean within-group effect sizes at follow-up (1.13), and the finding that

74% of treated children did not meet criteria for their primary diagnosis at the longest follow-up interval provide definite support for the maintenance of treatment effects.

One of the aims of the present meta-analysis was to establish the clinical significance of treatment effects. We aimed to answer the question, are treatment effects large enough to be clinically meaningful? (Lipsey & Wilson, 1993). Clinical significance findings for this meta-analysis, as measured using criteria developed by Jacobson and Truax (1991) and Kendall and Grove (1988; cited in Kendall, Marrs-Garcia, et al., 1999b), indicate that treatments produced moderately clinically significant outcomes. Thirty-eight percent of treatment conditions demonstrated clinically significant change at posttreatment (measured by RCI), and 54% of treatment groups were clinically equivalent to the normative group at posttreatment (as measured by normative comparison). Analyses of clinical significance at follow-up indicated the number of treatment conditions showing both a significant magnitude of change (58%) and clinical equivalency (83%) had increased from posttreatment. This suggests treatment gains were maintained and even increased over time.

3.2 Summary of Study Characteristics

In this meta-analysis we examined outcome research on the treatment of childhood anxiety disorders published between 1970 and 2001 was examined. This resulted in a total sample of 885 participants from 17 studies (plus two follow up studies). The majority of participants were recruited and treated from university based clinics rather than usual community settings. No studies included children below the age of 6 years and only one had a mean age greater than 12 years. The majority of participants were white and from middle class backgrounds and few were from minority groups. Studies tended to include participants with the more commonly diagnosed childhood anxiety disorders such as GAD, SAD, SOP, specific phobia and school phobia. Participants with less common childhood anxiety disorders, such agoraphobia, panic disorder and PTSD, were largely excluded.

Eighty-eight percent of the treatment conditions included in this current review were behavioural or cognitive-behavioural. Therefore it remains unclear

whether CBT was any more effective than other treatment approaches (Ollendick & King, 1998). There were no studies in this meta-analysis including psychodynamic, psychoanalytic, family systems or eclectic approaches. This is in keeping with the literature (e.g., Kazdin, 1994) as these approaches each account for less than 5% of techniques researched. In the process of including and excluding studies, it was noted that on occasions when published studies did employ these approaches, they tended to be single-case designs, lacked a comparison control or the sample did not meet full criteria for an anxiety disorder. Unfortunately this ruled such studies out of inclusion in this meta-analyses and highlight the need to increase controlled research of these approaches. This would also allow for a comparison with behavioural and cognitive-behavioural approaches, which was not possible in this meta-analysis.

Treatments were short and of limited duration and all used a manualised approach to treatment (Casey & Berman, 1985; Kazdin et al., 1990; Weisz et al., 1995). A number of studies used the Cognitive-Behavioural Program developed by Kendall and colleagues (1992), some used the 'Coping Cat Workbook' or a variation of it (e.g. Coping Koala, Coping Bear) (Barrett et al., 1996; Flannery-Schroeder, 2000; Kendall, 1994; Kendall et al., 1997; Mendlowitz et al., 1999; Shortt et al., 2001; Silverman, Kurtines, Ginsburg, Weems, Lumpkin, et al., 1999; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999). Alternatively, other studies incorporated the established principles and manuals of cognitive-behavioural paradigm and used their own manuals for specific diagnoses (e.g. Beidel et al., 2000; Hayward et al., 2000; Last et al., 1998; Öst et al., 2001; Spence et al., 2000).

Close to half of the studies (47%) included families in the treatment process. Comparison between the effect sizes of treatments including family (0.60) versus those that did not (0.76) indicated there was not a significant difference between treatments. The small number of studies included in this meta-analysis mean it is not possible to conclude whether or not the inclusion of family improves treatment outcomes. However, the results are consistent with conclusions in the literature indicating that family involvement may enhance treatment in the short-term but no significant difference has been found between treatments in the long-term (Barrett et al., 2001; Compton et al., 2002).

There were no significant findings between effect sizes of treatments conducted in group format (0.62) versus treatment delivered in individual format (0.77). While group mode treatments consistently produced positive outcomes there was no evidence to support the improved effectiveness of group interventions. Although group treatments were not shown to be superior to the individual format findings are important as they suggest group treatments have potential to be as effective as individual treatments. This may have clinical implications in terms of easing time and financial pressures in the delivery of treatment (Lumpkin et al., 2002).

Results also suggested that the effectiveness of treatments was similar across studies of different anxiety disorders. Effects were similar for studies including a homogeneous group for anxiety diagnoses and those including a range of anxiety disorders. This suggests treatment outcomes do not necessarily improve by restricting the range of anxiety disorders treated. Again as, Lumpkin and colleagues (2002) discuss this has implications for practice where the potential for treating a range of anxiety disorders may be desirable.

Treatments tended to be more effective when the sample consisted of more males than females. This result is in contrast to previous findings which have suggested that girls tend to show larger treatment responses than boys (Weisz et al., 1995). However, numbers were small in this meta-analysis and further information is needed before definite conclusions can be reached on the influence of gender on treatment outcome.

As well as improving symptoms of anxiety, outcomes suggested treatments were effective in improving behavioural avoidance, depression, social skills, coping and cognitive functioning. While numbers were small, mean effect sizes suggested behavioural avoidance (as measured by BATs) and coping functioning were associated with larger effects. This is consistent with Weiss and Weisz (1990) who found 'hard' outcomes (e.g., observed incidence of behaviour) were associated with larger effects. Furthermore, treatments focus on reducing avoidance and developing adaptive coping. Therefore these areas of functioning are expected to show improvement. Results also confirm previous findings that anxiety treatments do tend

to have a positive effect on depressive symptoms (Kendall, 1994). Studies reporting an outcome measure for depression produced a mean effect size of moderate magnitude (0.56). Conversely, anxiety focused treatments do not always target social functioning (Flannery-Schroeder & Kendall, 2000). Thus, it is not at all that surprising that measures of social skills were associated with smaller effects.

Findings suggested that different sources of assessment influenced outcomes, consistent with the findings of Weisz and colleagues (1995). Self-report and parental report outcomes were associated with effects of moderate magnitude (between 0.50 and 0.79), compared to teacher reports that produced very small effects (0.05). Although numbers were small in this group, this finding was similar to that of Weisz and colleagues (1995), and Kendall (1994) who proposed teacher reports may not be sensitive and thus fail to indicate treatment gains.

Analyses revealed an interaction between the type of control and magnitude of effect. Studies using a waitlist control produced an effect size in the large range (0.86), compared to those using an attention placebo control which resulted in an effect size with of a moderate magnitude (0.56). While the number of studies included in this meta-analysis was again not adequate to form sound conclusions, results do suggest comparison with an attention placebo control tends to reduce overall treatment effects (Kazdin, 1994).

Furthermore, within-group gain effect sizes indicated children in the attention placebo control conditions (0.50) showed considerably larger gain than those in waitlist control conditions (0.22). In two studies attention placebo control conditions demonstrated similar patterns of improvement to those in the cognitive-behavioural conditions in both the short and long term (Last et al., 1998; Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999a). Both studies used an Education Support intended to only provided support and non-specific therapeutic factors (e.g. education, therapist contact, warmth). Additionally, Silverman and colleagues demonstrated gains were maintained at 12-month follow-up with a within-group effect size similar to that of treatment conditions (1.21).

It is clear that the attention placebo control conditions in this review were better than no treatment at all. Again, while a low number of studies prevents firm

conclusions, results are consistent with current literature which suggest children with anxiety disorders may respond to a number of treatment approaches and the most active elements of treatment have yet to be identified (Compton et al., 2002).

Studies included in this meta-analysis were relatively well controlled in terms of design according to the methodological criteria. The inclusion of measures of treatment integrity and blind raters provided further testimony to the strength of designs. Additionally, the most well-controlled studies had lower rates of attrition and included an attention placebo control comparison. Although studies were divided in to two groups based on quality, studies in the lower group were in fact well controlled designs. For example, all but one study had sample sizes of more than 30, al randomly assigned participants, all included at least one normed outcome and all reported posttreatment data for all measures.

Contray to findings from previous meta-analyses, results suggested studies with lower scores for methodological quality (i.e. less methodologically sound) were associated with larger effects (0.81) than studies scoring a higher rating for quality (0.56). Only five studies obtained a rating of 5 or more and this was largely due to few meeting criteria for low attrition and an attention placebo control. As the numbers are small in this meta-analysis, it is not fully certain whether the quality of the studies in this review affected outcomes. However, while there is strong evidence in the meta-analysis literature that methodological rigor tends to be positively correlated with effect size (e.g. Hunter & Schmidt, 1990; Weisz & Hawley, 1998), the findings here suggest an oppose trend. The reason for this is not clear.

4.3 Methodological Issues

4.3.1 Limitations of Meta-analyses

The purpose of a meta-analysis is to provide a quantitative estimate of overall results from existing research of outcomes of treatment for childhood anxiety disorders. In doing this, it is also important to consider some of the methodological limitations which may have impacted on the findings.

While power analysis indicates the sample size for this meta-analysis was

sufficient to detect small, medium and large effects, the 17 included studies produced inadequate power in moderator analysis when studies are divided in to subgroups (Durlak et al., 1999; Hunter & Schmidt, 1990). Despite efforts to obtain missing data, moderator analysis was further hindered as a number of studies did not include information on variables such as mean age or the ethnicity of participants. The absence of this data and a small number of studies meant the influence of these variables on outcomes could not be evaluated.

A common criticism of meta-analyses is the bias created in failing to consider studies that have not been published. Omitting this research is thought to artificially increase the magnitude of effect as few, if any, studies supporting the null hypothesis are likely to be included (Glass et al., 1981). In this meta-analysis file draw analysis suggests approximately fifty unpublished studies would be needed to reduce the effect to a small magnitude. It is unlikely this many unpublished methodologically sound studies with null results exist.

It is possible relevant studies were overlooked. The way in which studies are identified for a meta-analysis is problematic. For example, research of online databases for the identification of relevant studies indicates their sensitivity may be less than adequate (Dickersin et al., 1994; Glass et al., 1981; Lipsey & Wilson, 2001). In this meta-analysis, a large number of studies were identified by online databases which were outside the inclusion criteria. Efforts were made to identify all relevant studies, including the searching of multiple databases, hand searching of major journals identified as likely to contain outcomes studies on the treatment of anxiety disorders in children, and contacting prominent researchers in the area.

A publication issue which could not be addressed in this meta-analysis pertains to the difficulty keeping up with new research (Michael and Crowley, 2002). All studies in this meta-analysis were published in the last 10 years which suggests this is a rapidly developing area. It is feasible that recently published important findings may not be included.

This meta-analysis also has limitations due to the selection criteria for studies. For example, selection criteria limited studies to those that diagnosed DSM anxiety disorders in all participants. It was noted that this excluded some otherwise

methodologically sound studies, as they included participants with sub-clinical anxiety or had not made a diagnosis with a clinical interview. This process also excluded a number of studies with participants with anxiety disorders not commonly found in child participants, such agoraphobia, panic disorder and PTSD (e.g. Deblinger et al., 1996; King, Tonge, Mullen, Myerson, Heyne, Rollings, Martin, & Ollendick, 2000; Krakow, Sandoval, Schrader, Keukne, McBride, Yau & Tandberg, 2001).

Eysenck (1994) raised a number of problems surrounding meta-analysis which are relevant to this review. The first involves the criticism of the early meta-analysis of psychotherapy outcomes conducted by Smith and Glass (1977; cited in Eysenck, 1994) for comparing patients, and treatments, which varied widely. It can be argued that the current meta-analysis is more homogeneous as the sample included only children with anxiety disorders. As previously discussed, the range of anxiety disorders included in the samples was also quite narrow. Treatments were also comparable as all treatments were either behavioural or cognitive-behavioural in nature.

Second, Eysenck (1994) criticised meta-analyses for including studies with varying degrees of methodological quality. While this is a concern, the inclusion criteria for this review selected studies with sound designs and overall methodological analysis revealed studies were generally well controlled.

4.3.2 Calculation of Effects

Effect sizes were calculated in this meta-analysis in a number of ways. These included between-group effect sizes, within-group effect sizes, and effect sizes calculated from proportions of treatment and control groups found to be diagnosis free. Multiple measures of effect were used to both utilise all available data and to allow for treatment outcomes to be described and evaluated comprehensively by multiple means. The result was that it was possible using multiple indexes of treatment outcomes, to establish that treated participants were significantly better off than control group participants and that treatment groups improved from pretreatment

to posttreatment and pretreatment to follow-up. Each of these methods is now discussed.

Between-group effect sizes were calculated using the traditional method comparing treatment and control group means at posttreatment or follow-up. The effect size index *g* was chosen as it adjusts for variance and sample size. The pooled standard deviation was used. While Hunter and Schmidt (1990) recommend the pooled standard deviation, there is some evidence that treatment and control group standard deviations can differ and it is therefore inappropriate to combine them (Weiss & Weisz, 1990; Weisz et al., 1995). The theory is that treatment increases variability and alternately the control is unaffected by treatment and is thus less variable (Glass et al., 1981; Lipsey and Wilson, 2001).

However, this suggestion assumes the control group to be untreated. In this meta-analysis the control groups in five studies received a form of support therapy and therefore may have been more variable than untreated control groups. The use of the pooled standard deviation could be considered a limitation of this meta-analysis and it may have been advisable to compare effect sizes calculated with a pooled standard deviation compared to those calculated with the control group standard deviation. Unfortunately this was not possible due to a software limitation where it was only possible to calculate effect sizes using the pooled standard deviation (M. Borenstein, personal communication, 7 January 2003). The implication of this may be that resulting effect sizes were conservative.

The method used to pool effect sizes was to average effect sizes within studies. This method was used both in the calculation of the overall effect size and in calculating a mean effect size for each study. As mentioned earlier, when this method was compared with other methods, the overall results suggested that different methods produced similar results. Nevertheless, it is maintained that the method chosen was the best choice for this meta-analysis as it avoids bias in selecting a particular effect size as all effect sizes contribute to the overall result (also see Weiss & Weisz, 1990).

When effect sizes are averaged within studies, a problem arises of statistical non-independence. Effect sizes are dependent as they are drawn from the same

sample. To address the problem of non-independence, it is necessary to calculate estimates of covariance between dependent effect sizes (Lipsey & Wilson, 2001). As this involves complicated analyses, this study in keeping with benchmark studies, did not go so far as to do this. Therefore non-independence of effect sizes remains an issue in the calculation of effect sizes in this meta-analysis and may cause an inflation of overall results (Hunter & Schmidt, 1990; Weiss & Weisz, 1990). Additionally the issue is further complicated when there is more than one treatment group and the control group is compared to multiple groups. In this meta-analysis we halved the sample size of the control group in studies where there was more than one treatment group to address the problem of multiple comparisons (Meyer & Mark, 1995).

As outlined previously, overall mean effect sizes were calculated from all measures of anxiety symptoms only. Effects were calculated for other constructs and analysed separately. The analysis provided support for our decision to include only measures of anxiety in the calculation of the overall mean effect size as constructs produced differential outcomes. If all measures had been pooled this may have resulted in an overly conservative estimate of the effect of treatments for anxiety disorders.

Including all anxiety outcome measures in the calculation of effect sizes introduces an issue around the reliability of measurement and the subsequent accuracy of effects. Hunter and Schmidt (1990) have suggested adjustments for differential reliability of measures should be made. However, consistent with benchmark studies, all outcomes were weighted equally. One recommendation to address this issue is to separate standardised and unstandardised measures and to calculate effect sizes for each (Kazdin et al., 1990). This could have been done in this meta-analysis. As suggested by Hunter and Schmidt (1990) standardised measures are more reliable and thus tend to be associated with larger magnitudes of effects. Therefore if including all measures in this meta-analysis confounded results it is most likely to have caused a conservative estimation of treatment effects.

Another issue which may have underestimated effects was the practice of estimating missing effects at zero. Fortunately this was only done for one study when calculating within-group effect sizes at follow-up. As Meyer and Marks (1995)

discuss this is an overly conservative method when, as in this meta-analysis, most effect sizes are non-zero and positive. Alternatively, this method addresses the problem of excluding a study altogether and thus reduces selection bias.

Finally, a random effects model was used in the analyses. While this seemed appropriate given the heterogeneity of effect sizes and increases external validity were goals, this presents two difficulties. First, as the random effects model includes an estimate of variability between as well as within studies, the resulting effects are smaller than the fixed effects model and confidence intervals are wider (Shadish, Matt, Navarro, Phillips, Ragsdale, & Heinsman, 2000). Therefore overall effects may have been underestimated. Second, it may not be appropriate to generalise beyond the included studies. As mentioned earlier, treatment outcome for childhood anxiety disorders is a rapidly expanding field, and it is possible, though perhaps not probable, emerging studies are not similar to the studies included in this current meta-analyses.

In addition to between-group effect sizes, we calculated within-group effect sizes to estimate the magnitude of change within conditions from pretreatment to posttreatment and pretreatment to follow-up. Within-group effect sizes are commonly criticised due to their lack of a control comparison. Interestingly, while the within-group mean effect size (pretreatment to posttreatment) was larger in magnitude it was not significantly larger than the between-group mean effect size (0.88 and 0.77 respectively). Within-group effect sizes were on average 14% larger than between-group effect sizes.

In this current meta-analysis the utility of within-group effect sizes was in estimating the magnitude of change at follow-up where few studies included a control comparison and thus between-group effect sizes could not be calculated. Within-group effect sizes for control groups were also useful in estimating change not due to treatment. The within-group effect sizes also indicated a potential difference in change between untreated control conditions and attention placebo control groups. Although not generally recommended within-groups effect sizes may have utility where comparison with a control may not be possible or change within a control condition is of interest, as in this meta-analysis.

Effect sizes were also calculated from proportions or percentages of participant's diagnosis free using probit transformations. While these effect sizes allow for the transformation of this data in to an effect size index, they are associated with a number of problems. First, the index is inflated as there are no adjustments available, including those for sample size (Cohen, 1988). The mean effect size based on proportions in this current meta-analysis was 44% larger than the between-group effect size. Second, this index of effect seemed to be of limited utility in providing useful information to readers. Stating 64% of treated participants compared to 15% of control participants seems to be a much more useful format than an overall mean effect size of 1.11.

As well as calculating various effect size indexes, analyses of clinical significance were also included in this meta-analysis to address questions regarding the meaningfulness of treatment outcomes. Two methods (RCI and normative comparison) were used to do this as recommended by Sheldrick and colleagues (2001). While these methods allowed for the demonstration of the clinical significance of treatments included this meta-analyses, there were some problems encountered. A notable difficulty relates to deciding on a normative group for comparison. When these analyses are conducted with individual data it is obvious which normative group to use. However, when using group data which includes a range of ages and both girls and boys, a decision needs to be made as to which normative group to use as a comparison. In this case, we used the most conservative normative group as recommended (Kendall, Marrs-Garcia, et al., 1999). Although not possible in this meta-analysis, individual data may give more accurate results (Sheldrick et al., 2001).

The issue of whether comparison with a community normative sample is appropriate has also been discussed in the literature (e.g. Kendall and Sheldrick, 2000; Sheldrick et al., 2001). These authors query whether comparing treatment participants with normal samples may be too rigorous criteria. They point out that treatments may represent clinically significant improvement even if participants are not equal to normative samples at posttreatment. Therefore some treatments which did not reach clinical significance in this current meta-analysis may still be

considered to have produced meaningful outcomes for children who began with significant problems, including reduced rates of diagnosis as a function of treatment.

A final shortcoming is that the evaluation of clinical significance was based only on one outcome measure, the CBCL-I. Additionally although the measure has acceptable reliability and validity it was potentially biased as it relies on parent report (Sheldrick et al., 2001). While it was not possible in this current meta-analysis, including a blind clinician rated outcome would have increased the strength of findings. As mentioned earlier, the RCMAS was not in the clinical range at pretreatment and thus had to be omitted from clinical significance analyses. This also meant studies that did not include this measure were omitted from the analysis. While this finding is difficult to explain, a possibility is that children in these samples under-reported their anxiety. As James, Reynolds and Dunbar (1994) discuss, self-report measures are reliant on children's willingness to disclose their fears and anxieties and in some cases they may be unwilling to admit to them.

Lastly, an aim of this meta-analysis was to compare between-group effect sizes calculated from gain scores with those calculated in the traditional way from post-treatment scores. Unfortunately we were not able to do this due to a lack of data for the calculation of between-group gain score effect sizes. As outlined earlier, the extensive steps made to contact authors and request the gain scores resulted in very little data becoming available. This highlights a difficulty in acquiring additional data from authors and the importance for gain scores to be included in published articles (Clark & Oxman, 2002). This is a significant limitation for this meta-analysis.

4.4 Recommendations for Future Research

This meta-analysis and the preceding literature review identified a number of gaps in the research and possible directions for the future. The current review identified only seventeen outcome studies addressing the treatment of children with anxiety disorders. Perhaps the most important recommendation is simply an increase in the amount of controlled research conducted on treatments for childhood anxiety disorders (Ollendick & King, 1998; Dadds et al., 1998).

As mentioned above, more longitudinal research is needed to establish whether treatments are producing enduring positive change (Kazdin, 1994; Ollendick & King, 1994). While 88% of the studies conducted follow up assessments, few included an assessment of the control group to allow a between-group effect size to be calculated at follow-up. As stated by a number of the included studies, the lack of comparison available at follow-up is often due to ethical concerns surrounding the with holding of treatment from controls (Beidel et al., 2000; Kendall, 1994; Silverman, Kurtines, Ginsburg, Weems, Lumpkin, et al., 1999).

Additional research is required to establish the efficacy of treatments including family members and group interventions and whether these additions to interventions lead to lasting improvements above other treatments (King & Ollendick, 1998). More studies are required to identify differential treatment responders to particular treatment modalities. For example, Heyne and colleagues (2002) have suggested family involvement may particularly benefit school refusing children who are younger and may be necessary when oppositional and defiant behaviours are also present. Manassis, Mendlowitz, Scapilato, Avery, Fiksenbaum et al. (2002), have also suggested children with higher levels of anxiety may respond better to individual treatment opposed to group treatments. Further research is required to establish whether such ideas are actually the case.

Prevalence studies have established that pre-school children also suffer from anxiety disorders (e.g. Richman, Stevenson, & Graham, 1975; cited in Bell-Dolan et al., 1990). No studies in this review included children younger than 6 years old in the sample. While it may be appropriate to restrict the age range of participants as questions have been raised about the differential treatment response of different age groups there has been little research targeting young children (Durlak et al., 1991; Weisz et al., 1995).

There has also been some indication that younger children may benefit more than older children from interventions including parents and due to their cognitive develop level they may find some components of individually focused CBT difficult (e.g. verbal reasoning) (e.g., Southam-Gerow et al., 2001). It is a concerning that there has not been adequate research to identify and evaluate treatments for pre-

school children with anxiety disorders. Future research is required to examine effective treatments for pre-school and younger school aged children which may require extra attention to child-parent relationships.

Chambless and Ollendick (2001) conclude that research on treatments for childhood anxiety disorders have shown specificity. That is some treatments are more or less effective with specific anxiety disorders. For example, exposure and modelling interventions are generally considered an efficacious treatment for simple phobias (Ollendick & King, 1998). However, they are less likely to be effective for GAD or panic disorder given the complexities of these disorders (e.g., Howard & Kendall, 1996; Ollendick, 1995). While this could not be investigated in the current meta-analysis, preliminary studies are emerging on treatments for children with the less commonly studied anxiety disorders, including panic disorder, agoraphobia, OCD and PTSD (Deblinger et al., 1996; Howard & Kendall, 1996; Ollendick, 1995; Thienemann et al., 2001; Waters et al., 2001). More controlled research is needed here to establish the effectiveness of treatments across anxiety disorders.

Although treatments have been shown to be acceptable to and effective for treatment completers, it is not clear that they are for those children who have failed to complete treatment (e.g. children from single parent homes and ethnic minorities). Attrition was an issue for a majority of the studies included in this review. Sixty-four percent of the studies had attrition greater than 10%. High attrition is problematic as it reduces sample size, undermining the power of a study and the conclusions that can be drawn from data (Kazdin, 1994). Kendall and Sugarman (1997) have discussed this issue in relation to the treatment of childhood anxiety disorders. In their study they examined the differences between treatment completers and terminators. One hundred and ninety children and their parents participated. Of the sample, 146 completed treatment and 44 did not. Findings indicated terminators were more likely to be from a single-parent household, be ethnic minorities, and report less anxious symptoms. Reflecting back to the risk factors associated with the development of anxiety disorders, it is concerning to note children at risk of developing an anxiety disorder are also those most likely to drop out of treatment.

Manassis and colleagues (2002) discuss a potential limitations of CBT treatments may be that they are less useful for families where regular attendance at sessions and completion of homework tasks may be difficult (e.g. impoverished families, single-parent families). This relates to the ethnic and demographic characteristics of study samples and is important in the New Zealand context. Currently it is unclear whether treatments would be appropriate in the New Zealand given the ethnic make-up of clients presenting at services. In this review, 28% of the sample were from ethnic minority groups. While this proportion seems reasonable, the percent was significantly inflated by the two studies conducted by Silverman and colleagues (1999a; 1999b) and Ginsburg and Drake (2002). These samples consisted of 37 % and 50% Hispanic children and 100% African American children. In total, only eleven of the seventeen studies actually reported data on ethnicity and only six reported socio-economic data. Improved reporting in studies of ethnicity and socio-economic status is required to allow the relationship with treatment outcomes to be evaluated.

Unfortunately, this meta-analysis does not answer questions about individual factors that may influence treatment outcome. Ferguson (2002) points out that group data and trends can not identify those individuals who may benefit from particular treatment approaches. Future research is needed to identify these moderators, and this may include regression analyses and hierarchical linear modelling (Ferguson, 2002).

Future research is also needed to address a number of methodological issues in treatment outcome studies for childhood anxiety disorders. Weisz and Hawley (1998) report more than 75% of studies included in meta-analysis to date are behavioural or cognitive behavioural in nature. This meta-analysis was no exception. This makes it difficult to develop a representative picture of the effectiveness of other treatment approaches. Future research would benefit from research looking at non-behavioural interventions, including psychodynamic, experiential, and family systems or ecological approaches.

Pharmacological treatments were not included in this review, but have been identified as potentially having a place in the treatment of childhood anxiety disorders (Dadds et al., 1998; Manassis, 2000; March, 1995). There is some research

supporting adjunctive psychopharmacotherapy with behavioural or cognitive-behavioural interventions for other conditions. Such as combination may be particularly appropriate for disorders such as OCD, panic disorder and children who have failed to respond to psychological treatments (Fairbanks, Pine, Tancer, Dummit, Kentgen, Martin, Asche, & Klein, 1997; March, 1995). Bernstein and colleagues (2000) have also suggested the use of adjunctive psychopharmacology in children with high levels of comorbid depression. Controlled outcome research is needed to clarify the efficacy of combining pharmacotherapy and CBT treatments for childhood anxiety disorders.

More studies are needed investigating attention placebo control conditions. Attention placebo conditions (i.e., non-directive therapies) have been shown to alleviate anxiety and further research on the non-specific elements of therapy is required to identify the key elements of treatments that contribute to positive treatment outcomes (Silverman, Kurtines, Ginsburg, Weems, Rabian, et al., 1999).

The inclusion of multiple parties and sources is recommended in published guidelines for the assessment and treatment of childhood anxiety disorders (Bernstein et al., 1997; NHC, 1998). Most of the studies in this review included family in the assessment process (94%). However, only three studies included teacher ratings and only one included teachers in treatment. This is interesting as four studies specifically addressed the treatment of school or social phobia and a further seven identified social phobia as one of the three primary diagnoses. As the school environment forms a significant part of a child's development, studies are recommended including teachers in interventions for anxiety disorders.

There is a great need for the research on childhood anxiety disorders to move from academic to naturalistic community settings. This is necessary to ensure treatments identified as efficacious in controlled outcome research on childhood anxiety disorders is transferable and effective for clients presenting in clinical practice (Lambert, 2001). Ginsburg and Drake's (2002) recent evaluation of a school-based programme with African-American youth is an important step in moving outcome research from academic settings to the community.

Finally, as discussed earlier the aim to compare two methods of calculating between-group effect sizes was not able to be done in this meta-analysis. Regrettably, while the utility of the between-group and within-group effect size indexes employed here was demonstrated, further research is needed to clarify whether there is a significant difference in overall outcomes from between-group effect sizes calculated from gain scores versus those calculated in the traditional way from post-treatment scores.

4.5 Implications for New Zealand

Evidence from this meta-analysis and relevant literature indicate that treatments reviewed here may be effective in treating childhood anxiety disorders in New Zealand. Studies examining the efficacy of behavioural and cognitive-behavioural interventions have been replicated and conducted in a number of different countries including Australia. It is likely that findings can also be generalised to the New Zealand context. Preliminary single case studies have indicated the success of cognitive-behavioural interventions with anxious children in New Zealand (Girling-Butcher & Ronan, 2003; Huzziff & Ronan, 2003)

In the introduction, anxiety disorders were identified as a significant problem for New Zealand children. New Zealand research has shown adolescents with anxiety disorders have an increased risk of experiencing anxiety, depression, substance dependence and academic underachievement later in life (Woodward & Fergusson, 2001). It is therefore important that services incorporate in to their practices research findings on efficacious treatments for childhood anxiety disorders. The literature and findings from this meta-analysis indicate behavioural and cognitive-behavioural treatments to be effective treatments. Consequently, these treatments should be examined more comprehensively in New Zealand community settings.

To date, guidelines such as those developed by the NHC (1998) have made recommendations regarding treatment. Cognitive-behavioural therapy implemented by competent professionals is a recommended treatment (NHC). In New Zealand, there are also resources available to guide practitioners in the process of assessing anxiety disorders in children (Ronan, 1996). This article links empirical evidence

with practices that can be implemented in clinical settings. In addition, treatment manuals (Girling-Butcher & Ronan, 2003; Huzziff & Ronan, 2003), have been developed that focus on family and behaviourally based variants of Kendall and colleagues programme, respectively. The first these (Girling-Butcher & Ronan) uses the idea of a "Coping Kiwi" as a metaphor for the development of an active coping template.

Although efficacious treatments have been identified and recommendations made, there may be barriers to their implementation. The Mental Health Commission (1998) has identified the lack of expertise in child and adolescent services as a barrier to the delivery of effective services to children and adolescents. Workforce development has been identified as a priority area in child and adolescent mental health services. This involves both the training of new staff and those currently employed in services.

Commitment to workforce development could lead to positive changes and improvements in the services provided to children with anxiety disorders. For example, Clinical Training Agency (CTA) has increased funding for child and adolescent mental health training and continues to fund training in cognitive-behaviour therapy (MoH, 1998). It is hoped training will provide the basis for the implementation and examination of the treatments reviewed in the current study.

Modifications to services may be necessary to complement New Zealand's cultural climate. As discussed in this review, the efficacy of treatments has not been adequately established for children from minority groups. It is vital modifications are made to ensure any interventions are acceptable to the New Zealand context. Additional single case studies may be useful initially in addressing this issue.

In conclusion, findings from this meta-analysis indicate childhood anxiety disorders can be treated effectively with behavioural and cognitive-behavioural interventions. Overall treatments produced positive effects in terms of the alleviation of anxiety symptoms and improvements in functioning. Follow up data suggests these benefits are maintained, or continue to improve, over time. New Zealand child and adolescent mental health services have been identified as deficient in terms of services and expertise. It is hoped knowledge of effective treatment strategies for

childhood anxiety disorders is considered in future workforce development and improvements to these services.

References

Alfano, C. A., Beidel, D. C., & Turner, S. M. (2002). Cognition in childhood anxiety: conceptual, methodological, and developmental issues. *Clinical Psychology Review, 22*(8), 1209-1238.

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research & Therapy, 31*(6), 621-631.

Ambrosini, P. J. (2000). Historical development and present status of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS). *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(1), 49-58.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th ed.).* Washington DC: American Psychiatric Association.

Anderson, J. C. (1994). Epidemiological issues. In D. G. Ollendick & N. J. King & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 43-65). New York: Plenum Press.

Anderson, J. C., Williams, S. M., McGee, R., & Silva, P. A. (1987). DSM-III disorders in preadolescent children: Prevalence in a large sample from the general population. *Archives of General Psychiatry, 44*(1), 69-76.

Anstendig, K. D. (1999). Is selective mutism an anxiety disorder? Rethinking its DSM-IV classification. *Journal of Anxiety Disorders, 13*(4), 417-434.

Axelson, D. A., & Birmaher, B. (2001). Relations between anxiety and depressive disorders in childhood and adolescence. *Depression & Anxiety, 14*(2), 67-78.

*Barrett, P. M. (1998). Evaluation of cognitive-behavioral group treatments for childhood anxiety disorders. *Journal of Clinical Child Psychology, 27*(4), 459-468.

*Barrett, P. M., Dadds, M. R., & Rapee, R. M. (1996). Family treatment of childhood anxiety: A controlled trial. *Journal of Consulting & Clinical Psychology, 64*(2), 333-342.

*Barrett, P. M., Duffy, A. L., Dadds, M. R., & Rapee, R. M. (2001). Cognitive-behavioral treatment of anxiety disorders in children: Long-term (6-year) follow-up. *Journal of Consulting & Clinical Psychology, 69*(1), 135-141.

Barrett, P. M., Rapee, R. M., Dadds, M. M., & Ryan, S. M. (1996). Family enhancement of cognitive style in anxious and aggressive children. *Journal of Abnormal Child Psychology, 24*(2), 187-203.

Barrett, P. M., Turner, C. M., & Sonderegger, R. (2000). Childhood anxiety in ethnic families: Current status and future directions. *Behaviour Change, 17*(3), 113-123.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective.* New York: Basic Books.

*Beidel, D. C., Turner, S. M., & Morris, T. L. (2000). Behavioral treatment of childhood social phobia. *Journal of Consulting & Clinical Psychology, 68*(6), 1072-1080.

Bell-Dolan, D. J., Last, C. G., & Strauss, C. C. (1990). Symptoms of anxiety disorders in normal children. *Journal of the American Academy of Child & Adolescent Psychiatry, 29*(5), 759-765.

Benjamin, R. S., Costello, E. J., & Warren, S. (1990). Anxiety disorders in a pediatric sample. *Journal of Anxiety Disorders, 4*(293-316).

Bernstein, G. A., Borchardt, C. M., & Perwien, A. R. (1996). Anxiety disorders in children and adolescents: A review of the past 10 years. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*(9), 1110-1119.

Bernstein, G. A., & Kinlan, J. (1997). Summary of the practice parameters for the assessment and treatment of children and adolescents with anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(11), 1639-1641.

Bernstein, G. A., Shaw, K., Dunne, J. E., Ayres, W., Arnold, V., Benedek, E., Benson, R. S., Bernet, W., Bryant, E., Gross, R. L., Jaffe, S., King, R., Kinlan, J., Leonard, H., Licamele, W., McClellan, J., Zametkin, A., Sloan, L. E., Miles, C. M., Black, B., Cantwell, D., Sylvester, C., & Warren, S. (1997). Practice parameters for the assessment and treatment of children and adolescents with anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(10 SUPPL.), 69S-84S.

Blanchard, E. B. (1970). Relative contributions of modelling, informational influences, and physicial contact in extinction of phobic behavior. *Journal of Abnormal Psychology, 76*(55-61).

Borenstein, M., & Rothstein. (2001). Comprehensive meta-analysis. Englewood, NJ: Biostat, Inc.

Boyd, C. P., Kostanski, M., Gullone, E., Ollendick, T. H., & Shek, D. T. L. (2000). Prevalence of anxiety and depression in Australian adolescents: Comparisons with worldwide data. *Journal of Genetic Psychology, 161*(4), 479-492.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98*(2), 388-400.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685-716.

Clarke, M., & Oxman, A. D. (Eds.). Cochrane reviewers handbook 4.1.5 [updated April 2002]. In: The Cochrane Library, Issue 2, 2002. Oxford: Update Software. Updated Quarterly.

Cobham, V. E., Dadds, M. R., & Spence, S. H. (1998). The role of parental anxiety in the treatment of childhood anxiety. *Journal of Consulting & Clinical Psychology, 66*(6), 893-905.

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

Compton, S. N., Burns, B. J., Egger, H. L., & Robertson, E. (2002). Review of the evidence base for treatment of childhood psychopathology: Internalising disorders. *Journal of Consulting & Clinical Psychology, 70*(6), 1240-1266.

*Cornwall, E., Spence, S. H., & Schotte, D. (1996). The effectiveness of emotive imagery in the treatment of darkness phobia in children. *Behaviour Change, 13*(4), 223-229.

Costello, E. J., Costello, A. J., Edelbrock, C., Burns, B. J., Dulcan, M. K., Brent, D., & Janiszewski, S. (1988). Psychiatric disorders in pediatric primary care. Prevalence and risk factors. *Archives of General Psychiatry., 45*(12), 1107-1116.

Craske, M. G. (1997). Fear and anxiety in children and adolescents. *Bulletin of the Menninger Clinic, 61*(2), 4-36.

Culbertson, J. L. (1999). Focus chapter: Research methods with children, *Kendall, Philip C. (Ed); Butcher, James N. (Ed); et al. (1999). Handbook of research methods in clinical psychology (2nd ed.).* (pp. 619-633). New York, NY, US: John Wiley & Sons, Inc.

Cummins, T. K., & Ninan, P. T. (2002). The neurobiology of anxiety in children and adolescents. *International Review of Psychiatry, 14*(2), 114-128.

Dadds, M. R., & Barrett, P. M. (2001). Psychological management of anxiety disorders in childhood. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 42*(8), 999-1011.

Dadds, M. R., Barrett, P. M., & Cobham, V. E. (1998). Anxiety disorders. *Comprehensive Clinical Psychology, 5.* Retrieved August 6, 2001, from the World Wide Web: http://www.sciencedirect.com/ccp

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 & Clinical Psychology, 67*(1), 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 & Clinical Psychology, 65*(4) Aug 1997), 627-635.

Davidson, P. R., & Parker, K. C. H. (2001). Eye movement desensitization and reprocessing (EMDR): A meta-analysis. *Journal of Consulting & Clinical Psychology, 69*(2), 305-316.

Deblinger, E., Lippmann, J., & Steer, R. (1996). Sexually abused children suffering posttraumatic stress symptoms: initial treatment outcome findings. *Child Maltreatment, 1*(4), 310-321.

Dickersin, K., Scherer, R., & Lefebvre, C. (1994). Identifying relevant studies for systematic reviews. *British Medical Journal, 309*, 1286-1291.

Dierker, L. C., Albano, A. M., Clarke, G. N., Heimberg, R. G., Kendall, P. C., Merikangas, K. R., Lewinsohn, P. M., Offord, D. R., Kessler, R., & Kupfer, D. J. (2001). Screening for anxiety and depression in early adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(8), 929-936.

Duggan, L., Fenton, M., Dardennes, R. M., El-Dosoky, A., & Idran, S. (2002). Olazapine for schizophrenia (Cochrane Review), *In: The Cochrane Library, Issue 1.* Oxford: Update Software.

Durlak, J. A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin, 110*(2), 204-214.

Essau, C. A., Conradt, J., & Petermann, F. (2002). Course and outcome of anxiety disorders in adolescents. *Journal of Anxiety Disorders, 16*(1), 67-81.

Eysenck, H. J. (1994). Meta-analysis and its problems. *British Medical Journal, 309*, 789-792.

Ferguson, C., J. (2002). Externalizing disorders and the treatment of child anxiety: A preliminary study. *Child & Behavior Therapy, 24*(3), 53-72.

*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 & Research, 24*(3), 251-278.

Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting & Clinical Psychology, 67*, 308-319.

Gilbody, S. M., & Petticrew, M. (1999). Rational decision-making in mental health: The role of systematic reviews. *The Journal of Mental Health Policy and Economics, 2*, 99-106.

Giles, T. R. (1990). Bias against behavior therapy in outcomes reviews: Who speaks for the patient? *The Behavior Therapist, 13*, 86-90.

*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 & Adolescent Psychiatry, 41*(7), 768-775.

Girling-Butcher, R., & Ronan, K. R. (2003). CBT and increased parental involvement for anxiety in youth: A multiple baseline study. Unpublished manuscript.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage Publications.

Graziano, A. M., & Mooney, K. C. (1980). Family self-control instruction for children's nighttime fear reduction. *Journal of Consulting & Clinical Psychology, 48*, 206-213.

Green, B. F., & Hall, J. A. (1984). Quantitative methods for literature reviews. *Annual Review of Psychology, 35*, 37-53.

Gullone, E. (2000). The development of normal fear: A century of research. *Clinical Psychology Review, 20*(4), 429-451.

Hampe, E., Noble, H., Miller, L. C., & Barrett, C. L. (1973). Phobic children one and two years posttreatment. *Journal of Abnormal Psychology, 82*, 446-453.

*Hayward, C., Varady, S., Albano, A. M., Thienemann, M., Henderson, L., & Schatzberg, A. F. (2000). Cognitive-behavioral group therapy for social phobia in female adolescents: Results of a pilot study. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(6), 721-726.

Health Funding Authority (2000). *Tuutahitia te wero, meeting the challenges: Mental health workforce development plan 2000-2005*. Christchurch, NZ: Health Funding Authority.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego: Academic Press, Inc.

Herbert, M. (1994). Etiological considerations. In D. G. Ollendick & N. J. King & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 3-20). New York: Plenum Press.

Heyne, D., King, N. J., Tonge, B. J., Rollings, S., Young, D., Pritchard, M., & Ollendick, D. G. (2002). Evaluation of child therapy and caregiver training in the treatment of school refusal. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(6), 687-695.

Hirshfeld, D. R., & Biederman, J. (2002). Rationale and principles for early intervention with young children at risk for anxiety disorders. *Clinical Child and Family Psychology Review, 5*(3), 161-172.

Howard, B. L., & Kendall, P. C. (1996). Cognitive-behavioral family therapy for anxiety-disordered children: A multiple-baseline evaluation. *Cognitive Therapy & Research, 20*(5), 423-443.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods in meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage Publications, Inc.

Huzziff, C. A., & Ronan, K. R. (2003). Exposure-based intervention for anxiety disordered youth: A multiple baseline study. Unpublished manuscript.

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting & Clinical Psychology, 67*(3), 300-307.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting & Clinical Psychology, 59*, 12-19.

James, E. M., Reynolds, C. R., & Dunbar, J. (1994). Self-report instruments. In D. G. Ollendick & N. J. King & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 317-329). New York: Plenum Press.

Kane, M. T., & Kendall, P. C. (1989). Anxiety disorders in children: A multiple-baseline evaluation of a cognitive-behavioral treatment. *Behavior Therapy, 20*(4), 499-508.

Kaplan, H. I., & Saddock, B. J. (1998). *Synopsis of psychiatry (8th ed.)*. Baltimore: Williams & Wilkins.

Kashani, J. H., & Orvaschel, H. (1990). A community study of anxiety in children and adolescents. *American Journal of Psychiatry, 147*(3), 313-318.

Kavale, K. A., & Forness, S. R. (2000). Policy decisions in special education: The role of meta-analysis, *Gersten, Russell Monroe (Ed); Schiller, Ellen P. (Ed); et al. (2000). Contemporary special education research: Syntheses of the knowledge base on critical instructional issues. The LEA series on special education and disability.* (pp. 281-326). Mahwah, NJ, US; Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc., Publishers; Lawrence Erlbaum Associates, Inc., Publishers.

Kazdin, A. E. (1994). Psychotherapy for children and adolescents. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change*. New York: John Willey & Sons.

Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting & Clinical Psychology, 58*(6), 729-740.

Keller, M. B., Lavori, P. W., Wunder, J., Beardslee, W. R., & et al. (1992). Chronic course of anxiety disorders in children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 31*(4), 595-599.

*Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting & Clinical Psychology, 62*(1), 100-110.

Kendall, P. C., Brady, E. U., & Verduin, T. L. (2001). Comorbidity in childhood anxiety disorders and treatment outcome. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(7), 787-794.

Kendall, P. C., Chansky, T. E., Kane, M. T., Kim, R. S., Kortlander, E., Ronan, K. R., Sessa, F. M., & Siqueland, L. (1992). *Anxiety disorders in youth: Cognitive-behavioral interventions*. Allyn & Bacon.

*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 & Clinical Psychology, 65*(3), 366-380.

Kendall, P. C., Flannery-Schroeder, E. C., & Ford, J. D. (1999). Therapy outcome research methods. In P. C. Kendall & J. N. Butcher & G. N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology (2nd ed.).* (pp. 330-363). New York: John Wiley & Sons, Inc.

Kendall, P. C., & Gosch, E. A. (1994). Cognitive-behavioral interventions. In D. G. Ollendick & N. J. King & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents. Issues in clinical child psychology.* (pp. 415-438). New York: Plenum Press.

Kendall, P. C., Kortlander, E., Chansky, T. E., & Brady, E. U. (1992). Comorbidity of anxiety and depression in youth: Treatment implications. *Journal of Consulting & Clinical Psychology, 60*(6), 869-880.

Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting & Clinical Psychology, 67*(3), 285-299.

*Kendall, P. C., & Southam-Gerow, M. A. (1996). Long-term follow-up of a cognitive-behavioral therapy for anxiety- disordered youth. *Journal of Consulting & Clinical Psychology, 64*(4), 724-730.

Kendall, P. C., & Sugarman, A. (1997). Attrition in the treatment of childhood anxiety disorders. *Journal of Consulting & Clinical Psychology, 65*(5), 883-888.

King, N. J., & Ollendick, T. H. (1997). Treatment of childhood phobias. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 38*(4), 389-400.

King, N. J., Ollier, K., Iacuone, R., Schuster, S., Bays, Gullone, E., & Ollendick, D. G. (1989). Fears of children and adolescents: A cross-sectional Australian study using the Revised-Fear Survey Schedule for Children. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 30*(5), 775-784.

King, N. J., Tonge, B. J., Mullen, P., Myerson, N., Heyne, D., Rollings, S., Martin, R., & Ollendick, T. H. (2000). Treating sexually abused children with posttraumatic stress symptoms: A randomized clinical trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(11), 1347-1355.

Kornhaber, R. C., & Schroeder, H. E. (1975). Importance of model similarity on extinction of avoidance behavior in children. *Journal of Consulting & Clinical Psychology, 43*(5), 601-607.

Kortlander, E., Kendall, P. C., & Panichelli-Mindel, S. M. (1997). Maternal expectations and attributions about coping in anxious children. *Journal of Anxiety Disorders, 11*(3), 297-315.

Krakow, B., Sandoval, D., Schrader, R., Keuhne, B., McBride, L., Yau, C. L., & Tandberg, D. (2001). Treatment of chronic nightmares in adjudicated adolescent girls in a residential facility. *Journal of Adolescent Health, 29*(2), 94-100.

Lambert, M. J. (2001). Psychotherapy outcome and quality improvement: Introduction to the special section on patient-focused research. *Journal of Consulting & Clinical Psychology, 69*(2), 147-149.

Last, C. G., Hansen, C., & Franco, N. (1997). Anxious children in adulthood: A prospective study of adjustment. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(5), 645-652.

*Last, C. G., Hansen, C., & Franco, N. (1998). Cognitive-behavioral treatment of school phobia. *Journal of the American Academy of Child & Adolescent Psychiatry, 37*(4), 404-411.

Leitenberg, H., & Callahan, E. J. (1973). Reinforced practice and reduction of different kinds of fears in adults and children. *Behaviour Research & Therapy, 11*(1), 19-30.

Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & et al. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III--R disorders in high school students. *Journal of Abnormal Psychology, 102*(1), 133-144.

Lewis, S. (1974). A comparison of behavior therapy techniques in the reduction of fearful avoidance behavior. *Behavior Therapy, 5*, 648-655.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*(12), 1181-1209.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications, Inc.

Lumpkin, P. W., Silverman, W. K., Weems, C. F., Markham, M. R., & Kurtines, W. M. (2002). Treating a heterogeneous set of anxiety disorders in youth with group cognitive behavioral therapy: A partially nonconcurrent multiple-baseline evaluation. *Behavior Therapy, 33*, 163-177.

Manassis, K. (2000). Childhood anxiety disorders: Lessons from the literature. *Canadian Journal of Psychiatry, 45*(8), 724-730.

Manassis, K., Mendlowitz, S. L., Scapillato, D., Avery, D., Fiksenbaum, L., Freire, M., Monga, S., & Owens, M. (2002). Group and individual cognitive-behavioral therapy for childhood anxiety disorders: A randomized trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(12), 1423-1430.

Manassis, K., & Menna, R. (1999). Depression in anxious children: possible factors in comorbidity. *Depression & Anxiety, 10*(1), 18-24.

March, J. S. (1995). Cognitive-behavioral psychotherapy for children and adolescents with OCD: A review and recommendations for treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 34*(1), 7-18.

March, J. S. (2002). Combining medication and psychosocial treatments: an evidence-based medicine approach. *International Review of Psychiatry, 14*(2), 155-163.

McGee, R., Feehan, M., Williams, S., & Anderson, J. (1992). DSM-III disorders from age 11 to age 15 years. *Journal of the American Academy of Child & Adolescent Psychiatry, 31*(1), 50-59.

*Mendlowitz, S. L., Manassis, K., Bradley, S., Scapillato, D., Miezitis, S., & Shaw, B. F. (1999). Cognitive-behavioral group treatments in childhood anxiety disorders: The role of parental involvement. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*(10), 1223-1229.

Mental Health Commission (1998). *Blueprint for mental health services in New Zealand.* Wellington, NZ: Mental Health Commission.

Menzies, R. G., & Clarke, J. C. (1993). A comparison of in vivo and vicarious exposure in the treatment of childhood water phobia. *Behaviour Research & Therapy, 31*(1), 9-15.

Messer, S. C., & Beidel, D. C. (1994). Psychosocial correlates of childhood anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 33*(7), 975-983.

Meyer, T. J., & Mark, M. M. (1995). Effects of psychosocial interventions with adult cancer patients: A meta-analysis of randomized experiments. *Health Psychology, 14*(2), 101-108.

Michael, K. D., & Crowley, S. L. (2002). How effective are treatments for child and adolescent depression? A meta-analytic review. *Clinical Psychology Review, 22*(2), 247-269.

Miller, L. C., Barrett, C. L., Hampe, E., & Noble, H. (1972). Comparison of reciprocal inhibition, psychotherapy, and waiting list control for phobic children. *Journal of Abnormal Psychology, 79*(3), 269-279.

Ministry of Health (1998). *Our children's health: Key findings on the health of New Zealand children*. Wellington, NZ: Ministry of Health.

Muris, P. (2002). Relationships between self-efficacy and symptoms of anxiety disorders and depression in a normal adolescent sample. *Personality & Individual Differences, 32*(2), 17-35.

Muris, P., Meesters, C., & Gobel, M. (2002). Cognitive coping versus emotional disclosure in the treatment of anxious children: A pilot-study. *Cognitive Behaviour Therapy, 31*(2), 59-67.

*Muris, P., Merckelbach, H., Holdrinet, I., & Sijsenaar, M. (1998). Treating phobic children: Effects of EMDR versus exposure. *Journal of Consulting & Clinical Psychology, 66*(1), 193-198.

Muris, P., Merckelbach, H., van Haaften, H., & Mayer, B. (1997). Eye movement desensitisation and reprocessing versus exposure in vivo. A single-session crossover study of spider-phobic children. *British Journal of Psychiatry, 171*, 82-86.

Murphy, C. M., & Bootzin, R. R. (1973). Active and passive participation in the contact desensitization of snake fear in children. *Behavior Therapy, 4*, 203-211.

National Health Committee (1998). *Guidelines for assessing and treating anxiety disorders*. Wellington, NZ: National Health Committee.

Obler, M., & Terwillier, R. F. (1970). Pilot study on the effectiveness of systematic desensitization with neurologically impaired children with phobic disorders. *Journal of Consulting & Clinical Psychology, 34*, 314-318.

*Öst, L.-G., Svensson, L., Hellstroem, K., & Lindwall, R. (2001). One-session treatment of specific phobias in youths: A randomized clinical trial. *Journal of Consulting & Clinical Psychology, 69*(5), 814-824.

Ollendick, T. H. (1979). Fear reduction in techniques with children. *Progress in Behavior Modification, 8*, 127-168.

Ollendick, T. H. (1995). Cognitive behavioral treatment of panic disorder with agoraphobia in adolescents: A multiple baseline design analysis. *Behavior Therapy, 26*(3), 517-531.

Ollendick, T. H., & King, N. J. (1994). Diagnosis, assessment, and treatment of internalizing problems in children: The role of longitudinal data. *Journal of Consulting & Clinical Psychology, 62*(5), 918-927.

Ollendick, T. H., & King, N. J. (1998). Empirically supported treatments for children with phobic and anxiety disorders: Current status. *Journal of Clinical Child Psychology, 27*(2), 156-167.

Ollendick, T. H., & Ollendick, D. G. (1997). General worry and anxiety in children. *In Session-Psychotherapy in Practice, 3*(1), 89-102.

Pfefferbaum, B. (1997). Posttraumatic stress disorder in children: A review of the past 10 years. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(11), 1503-1511.

Pine, D. S., Cohen, P., Gurley, D., Brook, J., & Ma, Y. (1998). The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Archives of General Psychiatry, 55*(1), 56-64.

Prendergast, M. L., Podus, D., Chang, E., & Urada, D. (2002). The effectiveness of drug abuse treatment: A meta-analysis of comparison group studies. *Drug & Alcohol Dependence, 67*(1), 53-72.

Rapee, R. M. (1997). Potential role of childrearing practices in the development of anxiety and depression. *Clinical Psychology Review, 17*, 47-67.

Rector, N. A., & Beck, A. T. (2001). Cognitive behavioral therapy for schizophrenia: An empirical review. *Journal of Nervous & Mental Disease, 189*(5), 278-287.

Reynolds, C. R., & Paget, K. D. (1983). National normative and reliability data for the revised children's manifest anxiety scale. *School Psychology Review, 12*(3), 324-336.

Ronan, K. R. (1996). Building a reasonable bridge in childhood anxiety assessment: A practitioner's resource guide. *Cognitive & Behavioral Practice, 3*(1), 63-90.

Ronan, K. R., Kendall, P. C., & Rowe, M. (1994). Negative affectivity in children: Development and validation of a self-statement questonnaire. *Cognitive Therapy & Research, 18*, 509-528.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638-641.

Rosenthal, R. (1994). Parametric measures of effect size, *Cooper, Harris (Ed); Hedges, Larry V. (Ed). (1994). The handbook of research synthesis.* (pp. 231-244). New York, NY, US: Russell Sage Foundation.

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

Sanders, M. R., Dadds, M. R., & Barrett, P. M. (1995). The prevention of anxiety and stress related disorders in children. In B. G. D. B. Raphael (Ed.), *Handbook of studies on preventive psychiatry* (pp. 399-421). New York: Elsevier Science, B.V.

Shadish, W. R., Matt, G. E., Navarro, A. M., Phillips, G., Ragsdale, K., & Heinsman, D. T. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin, 126*(4), 512-529.

Shadish, W. R., & Ragsdale, K. (1996). Random Versus Nonrandom Assignment in Controlled Experiments: Do You Get the Same Answer? *Journal of Consulting & Clinical Psychology, 64*(6), 1290-1305.

Shadish, W. R., & Ragsdale, K. (1998). Random Versus Nonrandom Assignment in Controlled Experiments: Do You Get the Same Answer? *Year Book of Psychiatry & Applied Mental Health Annual, 4*, 123-124.

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review, 17*(8), 881-901.

Sheldrick, R. C., Kendall, P. C., & Heimberg, R. G. (2001). The Clinical Significance of Treatments: A Comparison of Three Treatments for Conduct Disordered Children. *Clin. Psychol., 8*(4), 418-430.

Sheslow, D. V., Bondy, A. S., & Nelson, R. O. (1982). A comparison of graduated exposure, verbal coping skills, and their combination in the treatment of children's fear of the dark. *Child & Family Behavior Therapy, 4*(2-3), 33-45.

Shortt, A. L., Barrett, P. M., Dadds, M. R., & Fox, T. L. (2001). The influence of family and experimental context on cognition in anxious children. *Journal of Abnormal Child Psychology, 29*(6), 585-596.

*Shortt, A. L., Barrett, P. M., & Fox, T. L. (2001). Evaluating the FRIENDS Program: A cognitive-behavioral group treatment for anxious children and their parents. *Journal of Clinical Child Psychology, 30*(4), 525-535.

Silverman, W. K., Cerny, J. A., Nelles, W. B., & Burke, A. E. (1988). Behavior problems in children of parents with anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 27*(6), 779-784.

*Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Lumpkin, P. W., & Carmichael, D. H. (1999). Treating anxiety disorders in children with group cognitive-behaviorial therapy: A randomized clinical trial. *Journal of Consulting & Clinical Psychology, 67*(6), 995-1003.

*Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Rabian, B., & Serafini, L. T. (1999). Contingency management, self-control, and education support in the treatment of childhood phobic disorders: A randomized clinical trial. *Journal of Consulting & Clinical Psychology, 67*(5), 675-687.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*(9), 752-760.

Southam-Gerow, M. A., Kendall, P. C., & Weersing, V. R. (2001). Examining outcome variability: Correlates of treatment response in a child and adolescent anxiety clinic. *Journal of Clinical Child Psychology, 30*(3), 422-436.

*Spence, S. H., Donovan, C., & Brechman-Toussaint, M. (2000). The treatment of childhood social phobia: The effectiveness of a social skills training-based, cognitive-behavioural intervention, with and without parental involvement. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 41*(6), 713-726.

Target, M., & Fonagy, P. (1994). Efficacy of psychoanalysis for children with emotional disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*(3), 361-371.

Thieneman, M., Martin, J., Cregger, B., Thompson, H. B., & Dyer-Friedman, J. (2001). Manual-driven group cognitive-behavioral therapy for adolescents with obsessive-compulsive disorder: A pilot study. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(11), 1254-1260.

Ultee, C. A., Griffioen, D., & Schellekens, J. (1982). The reduction of anxiety in children: A comparison of the effects of "systematic desensitization in vitro" and "systematic desensitization in vivo." *Behaviour Research & Therapy, 20*(1), 61-67.

Warren, S., Huston, L., Egeland, B., & Sroufe, L. A. (1997). Child and adolescent anxiety disorders and early attachment. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 637-644.

Waters, T. L., Barrett, P. M., & March, J. S. (2001). Cognitive-behavioral family treatment of childhood obsessive-compulsive disorder: Preliminary findings. *American Journal of Psychotherapy, 55*(3), 372-387.

Weiss, B., & Weisz, J. R. (1990). The impact of methodological factors on child psychotherapy outcome research: A meta-analysis for researchers. *Journal of Abnormal Child Psychology, 18*(6), 639-670.

Weiss, B., & Weisz, J. R. (1995). Relative effectiveness of behavioral versus nonbehavioral child psychotherapy. *Journal of Consulting & Clinical Psychology, 63*(2), 317-320.

Weisz, J. R., & Hawley, K. M. (1998). Finding, evaluating, refining, and applying empirically supported treatments for children and adolescents. *Journal of Clinical Child Psychology, 27*(2), 206-216.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting & Clinical Psychology, 55*(4), 542-549.

Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & et al. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117*(3), 450-468.

Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting & Clinical Psychology, 69*(6), 875-899.

Wewetzer, C., Jans, T., Müller, B., Neudörfl, A., Bücherl, U., Remschmidt, H., Warnke, A., & Herpertz-Dahlmann, B. (2001). Long-term outcome and prognosis of obsessive-compulsive disorder with onset in childhood or adolescence. *European Child & Adolescent Psychiatry, 10*(1), 37-46.

Wolf, F. R. (1986). *Meta-analyis: Quantative methods for research synthesis.* Beverly Hills: Sage Publications.

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 & Adolescent Psychiatry, 40*(9), 1086-1093.

Appendix 1

Dear Dr [corresponding author]

My name is Kirsten McMurray I am a student enrolled in the clinical psychology programme at Massey University, New Zealand. For my Masters thesis I am undertaking a meta-analysis of treatment outcomes for children with anxiety disorders. This includes controlled studies of psychological interventions for children diagnosed with anxiety disorders. Dr Kevin Ronan is my supervisor for this thesis.

I am contacting you as we wish to include the following treatment study/s for which you are the corresponding author: [reference]

As well as establishing the effectiveness of psychological treatments for childhood anxiety disorders, an additional aim of this thesis concerns the methodology of meta-analysis. The aim is to compare two methods of calculating effect sizes: posttreatment scores (treatment group mean less control group mean) versus change scores (pretreatment to posttreatment change). Benchmark meta-analyses on the effectiveness of psychotherapy have tended to employ posttreatment scores (e.g., Casey & Berman, 1985; Weisz, Weiss, Alicke & Klortz, 1987; Weisz, Weiss, Han, Granger & Morton, 1995). However, more recent meta-analyses often include both methods of calculating effect sizes (e.g., Clarke & Oxman, 2002; Michael & Crowley, 2002; Rector & Beck, 2001). Currently, there has been very little written about the utility of these methods, whether there is likely to be a significant difference in effects, and in some instances it is unclear which method researchers have employed. In comparing these methods we hope to make recommendations on the best method of calculating effect sizes, considering the characteristics of data and what is most likely to be clinically meaningful and useful.

In order to include studies we require within-group change means and standard deviations (change from pretreatment to posttreatment or follow up) for all outcome measures, as well as the pretreatment and posttreatment means and standard deviations contained in your published article. This will enable us to calculate and

compare the two types of effect sizes.

From your published article we are able to calculate change means and could possibly make conversions to obtain some standard deviations. However, obtaining raw change score data from authors, will maximise accuracy in the calculations of effects and minimise the amount of missing data. This practice is in keeping with benchmark meta-analyses in the area of effects of psychotherapy (e.g., Michael & Crowley, 2001; Weisz et al., 1995) and is recommended in meta-analysis literature (e.g., Clarke & Oxman, 2002; Hunter & Schmidt, 1990).

We would appreciate it if you could indicate if you are willing to supply within-group change means and standard deviations for outcome measures in the aforementioned study/s.

I look forward to hearing from you.

Yours sincerely

Kirsten McMurray

References

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98*, 388-400.

Clarke, M., Oxman, A. D. (Editors). Special topics. *Cochrane Reviewers' Handbook 4.1.5* [updated April 2002]; Section 8. In: The Cochrane Library, Issue 2, 2002. Oxford: Update Software. Updated quarterly.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Thousand Oaks, CA, US: Sage Publications, Inc.

Michael, K. D., Crowley, S. L. (2002). How effective are treatments for child and adolescent depression? A meta-analytic review. *Clinical Psychology Review, 22* (2), 247-269.

Rector, N. A., & Beck, A. T. (2001). Cognitive behavioral therapy for schizophrenia: An empirical review. *The Journal of Nervous and Mental Disease, 189* (5),

278-287.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klortz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55*, 542-549.

Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117* (3), 450-468.

Appendix 2

Instruments

1. Children's Anxiety Sensitiviey Index (CASI; Silverman, Fleisig, Rabian, & Peterson, 1991)
2. Child Behavior Checklist (CBCL; Achenbach, 1991)
3. Children's Coping Strategies Checklist (CCSC; Program for Prevention Research, 1992)
4. Children's Depression Inventory (CDI; Kovacs, 1992)
5. Children's Negative Cognition Error Questionnaire (CNCEQ; Leitenberg, Yost, & Carroll-Wilson, 1986)
6. Coping Questionnaire – child version (CQ-C; Kendall, 1994)
7. Darkness Fear Behaviour Questionnaire (Cornwall, Spence, & Schotte)
8. Eysenck Personality Inventory (EPI; Eysenck & Eysenck, 1968)
9. Family Enhancement of Avoidant Responses (FEAR; Barrett, Rapee, Dadds, & Ryan, 1996)
10. Fear Thermometer (FT; Melamed & Siegel, 1975)
11. Friendship Measure – Child (FM-C; Bierman & McCauley, 1987)
12. Fear Survey Schedule for Children – Revised (FSSC-R; Ollendick, Matson, & Helsel, 1985)
13. Global Improvement Scale (Gittelman-Klein & Klein, 1971)
14. Global Assessment Scale (K-GAS; ShafferGOuld, Brasie, Abrosini, Fisher, Bird, & Aluwahlia, 1983)
15. Loneliness Scale (LS; Asher, Hymel, & Renshaw, 1984)
16. Children's Negative Affectivity Self-Statement Questionnaire (NASSQ; Ronan, Kendall, & Rowe, 1994)
17. Parent's Rating of Child's Competence (PRSC; Harter, 1982)
18. Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1985)
19. Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher, Khetarpal, Brent, & Cully (1997)
20. Self Assessment Manikin (SAM; Hodes, Cook, & Lang, 1985)
21. Social Anxiety Scale for Children –Revised (SASC-R; La Greca & Stone, 1993)
22. Social Activities Scale - Parent (SAS-P; Flannery-Schroeder & Kendall, 2000)
23. Spence Children's Anxiety Scale (SCAS; Spence, 1997)
24. Social Competence Quesionnaire - Parent (SCQ-P; Spence, 1995)
25. Social Phobia and Anxiety Inventory (SPAI; Turner, Stanley, Beidel, & Bond, 1989)
26. Self Perception Profile for Children (SPPC; Harter, 1985)
27. Social Skills Questionnaire – Parent (SSQ-P; Spence, 1995)
28. State-Trait Anxiety Inventory for Children (STAIC; Spielberger, 1973)
29. Social Worries Questionnaire – Pupil (SWQ-PU; Spence, 1995)

Ultimate Outcomes

30. Diagnosis - Anxiety Disorders Interview Schedule for Children, Diagnosis (ADIS-C; Silverman & Albano, 1997); Diagnostic Interview Schedule for Children, Adolescents and Parents, Diagnosis (DISCAP; Holland & Dadds, 1995); Schedule for Affect Disorders and Schizophrenia for School-Age Children (K-SADS) (Chambers, Puig-Antich, Hirsh, Paez, Ambrosini, Tabrizi & Davies, 1985)
31. ADIS Severity ratings
32. ADIS Clinician rating (Global, Anxiety, Avoidance, Child Skill, Family, Family Skill, Family Functioning)
33. Negative coping (% of events)
34. School Attendance
35. Stressful events (% of days)

Observational Outcomes

36. Behavioral Assertiveness Test for Children (BAT-CR; Ollendick, 1981)
37. BAT Diastolic BP
38. BAT Heart rate
39. BAT Self rated Anxiety
40. BAT Systolic BP
41. BAT State Anxiety
42. Absence of eye contact
43. Fingers in mouth
44. Trembling voice
45. Darkness tolerance test
46. Read-aloud anxiety
47. Read-aloud effectiveness
48. Read-aloud self-rating
49. Role play anxiety
50. Role play effectiveness
51. Role-play self-rating