EDUCATION AND TREATMENT OF CHILDREN Vol. 37, No. 2, 2014

A Meta-Analytic Evaluation of the FRIENDS Program for Preventing Anxiety in Student Populations

Daniel M. Maggin

University of Illinois at Chicago

Austin H. Johnson

University of Connecticut

Abstract

The purpose of this review was to evaluate the methodological strength and overall effectiveness of the research underlying the FRIENDS program for preventing anxiety in students at low and elevated risk for developing anxiety disorders. Meta-analytic findings provided mixed results, with low-risk students exposed to the program having demonstrated small improvements over comparisons for immediate posttest measures of anxiety. Findings drawn from follow-up data collection periods indicated that low-risk students sustained initial gains on anxiety over 12 months but not beyond. In addition, no immediate posttest difference was found between students at elevated risk on measures of anxiety. These findings are discussed in terms of practical and methodological limitations of the body of research.

KEYWORDS: Anxiety, FRIENDS, Meta-analysis, Program Evaluation

The presence of anxiety in student populations poses significant challenges for school psychologists and other school personnel (Tomb & Hunter, 2004). For instance, traditional school-based assessment methods are often unable to distinguish reliably between typical and atypical levels of worry, fear, and apprehension often associated with anxiety disorders (Levitt, Saka, Romanelli, & Hoagwood, 2007). Not only have the measures for identifying students with anxiety been shown to be unreliable (Nemeroff et al., 2008), but students with palpable anxieties often avoid anxiety provoking circumstances, making the disorder difficult to observe and assess (Layne, Bernat, Victor, & Bernstein, 2008). These impediments to accurate detection have led to chronic under identification of anxiety disorders in students (Collins, Westra, Dozois, & Burns, 2004).

Although students with anxiety problems comprise only a small fraction of those receiving services for behavioral problems in

Correspondence to Daniel M. Maggin, Department of Special Education, University of Illinois at Chicago, 1040 West Harrison Street, M/C 147, Chicago, Il 60607; e-mail: dmaggin@uic.edu

Pages 277–306

schools, epidemiological studies indicate such disorders may be the most prevalent of all psychiatric conditions in school-aged populations (Beesdo, Knappe, & Pine, 2009; Davis, May, & Whiting, 2011). Moreover, these prevalence rates likely represent an underestimate of the pervasiveness of anxiety problems in students, because they refer only to individuals who demonstrate chronic levels of clinical anxiety (Cartwright-Hatton, McNicol, & Doubleday, 2006). In fact, research has indicated that nearly 30% of students will exceed clinical cutoff scores on self-reported indices of anxiety at some point during childhood or adolescence (Kessler et al., 2005).

The incidence rates of anxiety are concerning given that their onset is associated with greater risk for both academic and social failure (Duchesne, Vitaro, Larose, & Tremblay, 2008; Mychailyszyn, Mendez, & Kendall, 2010; Rubin, Burgess, Kennedy, & Stewart, 2003; Tomblin, Zhang, Buckwalter, & Catts, 2000). The inability to successfully interact with peers may also lead students with anxiety to be socially rejected (Rubin, Root, & Bowker, 2010). Taken together, the academic problems and social isolation associated with anxiety increases student risk for experiencing deleterious outcomes such as school dropout, substance abuse, and unemployment (Vaughn et al., 2010) as well as developing behavioral manifestations such as suicidal tendencies, eating disorders, and sexual promiscuity (Kauffman & Landrum, 2009).

These deleterious effects of anxiety have led to the development of a variety of intervention frameworks and methods (McLoone, Hudson, & Rapee, 2006). However, the procedures with greatest empirical support are typically provided to students only after a problem has been diagnosed rather than used preventatively (Davis et al., 2011). Despite the presence of these intensive, individualized treatment options, several researchers have focused on developing preventative interventions for anxiety that would be delivered within school classrooms (Neil & Christensen, 2009). These prevention-based strategies represent a proactive method for assisting students in building coping strategies and resiliency to manage their anxieties. Such methods are particularly important for those students who remain unidentified or experience notably high, though nonclinical, levels of anxiety. Recent reviews on school-based prevention programs for anxiety have shown these approaches to be generally effective (Horowitz & Garber, 2006; Schoenfeld & Janney, 2008). Despite these positive findings, there remain questions regarding the generality of these conclusions to individual programs. That is, programs are developed with a specific set of procedures that are intended to be implemented with integrity to increase the likelihood that the magnitude of treatment effects observed in practice are consistent with those obtained

during clinical trials (Shadish, Cook, & Campbell, 2002). This natural variation of components across different programs makes it critical to consider the research underlying individual programs in isolation rather than assume that a broad class of interventions, such as those designed to prevent anxiety, are generally effective. The purpose of the present review, therefore, was to evaluate the methodological strength and overall treatment effectiveness for a single, widely endorsed program delivered at the school and classroom levels.

The FRIENDS for Life Program

The FRIENDS for Life program is a commercially marketed curriculum that was developed to provide school personnel with a universal approach for preventing the onset of anxiety in school-aged populations (FRIENDS; Barrett, Lowry-Webster, & Turner, 2000; Barrett & May, 2007). Based on the Coping Cat program (Kendall, 1994), FRIENDS utilizes the principles of cognitive-behavioral therapy to explicitly teach students strategies for managing symptoms associated with generalized anxiety (Barrett & Pahl, 2006). These strategies are aligned with the FRIENDS acronym: *F* – Feeling worried?; *R* – Relax and feel good; *I* – Inner thoughts; *E* – Explore plans; *N* – Nice work so reward yourself; *D* – Don't forget to practice; and *S* – Stay calm, you know how to cope. In accordance with cognitive-behavioral theory, program components emphasize the development of emotional resilience and physiological awareness and are delivered over the course of 10 weekly 60-min sessions. Program content is presented to students via a series of structured activities such as role-playing, peer coaching, and class discussions. To review program material and concepts, a booster session is provided to students after the first and third months following program completion. In addition, parents typically are offered two days of comprehensive training to increase their awareness about the development and signs of anxiety disorders and to learn about program content.

The FRIENDS program has been the subject of nearly 50 studies since its development in 1998 with a majority of these being conducted in clinical settings. Despite the relatively limited research of FRIENDS in schools, the World Health Organization (WHO; 2004) identified it as the only school-based anxiety prevention program with sufficient empirical support to warrant classification as an evidence-based practice. A series of recently published systematic reviews have also indicated that the FRIENDS program is generally effective in reducing the onset of anxiety in children and youth (Horowitz & Garber, 2006; Neil & Christensen, 2009; Schoenfeld & Janney, 2008; Stallard, 2010). Unfortunately, these claims are based on research that has both

conceptual and methodological limitations including, (a) the use of narrative, rather than systematic review procedures (Schoenfeld & Janney, 2008; Stallard, 2010), (b) the inclusion of FRIENDS evaluations that were not conducted in school settings (Schoenfeld & Janney, 2008; Stallard, 2010), (c) the failure to consider the importance of methodological quality of included studies (Stallard, 2010; WHO, 2004), and (d) the focus of reviews on the broader prevention literature related to anxiety rather than the FRIENDS program exclusively (Horowitz & Garber, 2006; Neil & Christensen, 2009). Given that FRIENDS has been internationally recognized and commercially marketed as a school-based prevention program, it is important to provide a critical evaluation of the literature pertaining to versions of the FRIENDS curriculum used in schools and classrooms. This review was conducted, therefore, to determine whether the research underlying *school-based* versions of the FRIENDS program is sufficiently rigorous to support the assertion that FRIENDS is an evidence-based practice. A brief introduction to evidence-based practice and the criteria used to evaluate empirical strength is provided below.

Criteria for Identifying Evidence-Based Practices

Evidence-based practice refers to the careful consideration of research to augment clinical expertise when making programming decisions (Kratochwill & Shernoff, 2004). Following the identification of a clinical problem, such as the need to provide students with coping strategies to protect against the development of anxiety, school personnel are encouraged to identify programs and interventions most likely to produce the desired outcomes. The potential effectiveness of a given program is established through a series of rigorously conducted research studies in which participants exposed to the intervention have consistently demonstrated positive results when contrasted with individuals enrolled in a concurrent control condition. These studies subsequently serve as the basis for systematic reviews that provide comprehensive evaluations of the quality of the research methods and the robustness of the research findings (Cook, Mulrow, & Haynes, 1997). When evaluating the quality of scientific treatment research, the general factors considered include (a) the quality of reporting on sample characteristics and intervention features to determine for whom a particular treatment is most likely to produce the anticipated outcomes; (b) the extent to which rigorous research methods were used to minimize bias within individual studies including the use of random assignment, outcome measures which yield reliable scores, and appropriate statistical analyses; (c) the total number of studies investigating the program and the size of the samples used

to detect meaningful effects; and (d) the magnitude and consistency of the effects observed (e.g., Moher, Liberati, Tetzlaff, & Altman, 2009). Within the current review, a conscious effort was made to identify and describe relevant sources of error that might undermine the meaningfulness of observed effects to provide a clearly justified assessment of the potential effectiveness of the FRIENDS program.

Synthesis Purpose and Guiding Research Questions

The purpose of the present meta-analysis was to evaluate the methodological strength and overall treatment effectiveness of school-based versions of the FRIENDS program for preventing the incidence of anxiety in student populations. Specifically, the efficacy of the FRIENDS program was examined for students at both low and elevated risk for developing anxiety disorders to determine whether the program was effective for preventing the onset of anxiety for typically developing students as well as those students with the greatest need. As such, the following five research questions guided the evaluation: (1) What are the sample and program characteristics of studies examining school-based applications of the FRIENDS program? (2) What are the methodological strengths and weaknesses of school-based evaluations of the FRIENDS program? (3) How effective was the FRIENDS program for inhibiting the onset of anxiety in students at low risk for developing anxiety disorders? (4) How effective was the FRIENDS program for inhibiting anxiety in students at elevated risk for developing anxiety disorders? (5) How did students enrolled in treatment classrooms compare to those enrolled in control classroom at subsequent follow-up data collection periods on measures of anxiety?

Method

Article Identification Procedures

A three-step process was used to identify candidate studies for inclusion in the present review. First, candidate studies were located using an electronic search of databases that included literature that is both commercially controlled (e.g., book chapters, journal articles) and commercially unavailable (e.g., conference papers, technical reports). The specific databases searched included: Dissertation Abstracts, Educational Resources Information Center (ERIC), Google Scholar, Medline, PsychEXTRA, PsychINFO, PubMED, Scopus, Sociological Abstracts, and Thomson Reuters' Web of Knowledge. In order to ensure that only the most relevant studies were returned, searches were performed using the formal program titles (i.e., "FRIENDS," "Friends for Life," "Friends for Children," and "Friends

for Youth") and a delimiter corresponding to the year of program development (i.e., 1998). The title and abstract of each study from the database search was reviewed to determine if the article should be read in greater detail. Candidate studies were selected for more thorough review if the paper was deemed as having potential for providing an experimental or quasi-experimental evaluation of the FRIENDS program. Second, an ancestral search was conducted using the reference lists of each potentially eligible study located via the database search. Ancestral searches were used to identify additional studies that were not captured in the initial search of databases. Third, the citation lists of review articles and program materials were consulted to identify any remaining studies that might qualify (e.g., Horowitz & Garber, 2006; Neil & Christensen, 2009; Schoenfeld & Janney, 2008; Stallard, 2010).

Inclusion Criteria

The following six criteria were used to determine if particular studies were eligible for inclusion. First, the study had to utilize the FRIENDS program or one of its variations such as "Friends for Youth" developed by Barrett and colleagues (Barrett et al., 2000). Second, the study had to be conducted with students enrolled in kindergarten to grade 12 classrooms. Third, the study had to be conducted in a school or classroom environment. Consequently, those studies that evaluated the effectiveness of FRIENDS within a clinical setting were not included in the present review. Fourth, a standard measure of anxiety with previously demonstrated psychometric properties had to be used to assess treatment outcomes. The use of standard measures increases the reliability of research findings and ensures comparability across assessments. Fifth, a group-based experimental or quasi-experimental design with a control group had to be employed to evaluate the effectiveness of the FRIENDS program. Although single-subject studies may be able to provide a rigorous assessment of the relation between independent and dependent variables, single-case studies were not included in the present synthesis given the current challenges of combining effect sizes from group and single-case research designs (Shadish & Rindskopf, 2007). Sixth, the article had to be published in English.

Following these procedures, a total of 1,258 abstracts were examined with 1,207 being screened out upon initial review. Efforts were made to retrieve the studies associated with these 51 abstracts with 2 being unable to be collected. It should be noted that the two studies that were unable to be obtained were from conference presentations. A subsequent review of the remaining 49 studies resulted

in the further paring of eligible studies to 17 which were drawn from 16 manuscripts. Reasons for exclusion included (a) the use of an independent variable that was not the FRIENDS program or one of its variants ($n = 11$); (b) the study took place in an ineligible setting such as a clinic or community-based setting ($n = 9$); (c) the study utilized an ineligible evaluation method such as single-case design or non-experimental methods ($n = 4$); (d) the study was not published in English ($n = 1$); and (e) the study was not an intervention study ($n = 7$).

Study Coding Procedures

A goal of meta-analytic reviews is to aggregate research findings across several studies to determine the populations for whom and circumstances under which treatment effects are maximized (Lipsey & Wilson, 2001). To help synthesize the findings from the 17 eligible studies, a coding system was developed to describe various aspects of included studies. All items were drawn or adapted from the American Psychological Association's Task Force on Evidence-Based Interventions in School Psychology coding manuals (Kratochwill et al., 2002). The items of the present coding system related to (a) participant and setting characteristics, (b) program features, and (c) methodological features. The coding system for these items required the observer to report whether or not the feature was reported and, if present, to record the corresponding value for both the treatment and control group if applicable. A description of each of these domains is provided below.

Participant and setting characteristics. A series of items was developed to determine the specific demographic and setting characteristics of the students included across eligible studies. Items included (a) the percentage of male participants, (b) the mean age of the sample, (c) the mean grade of the sample, (d) the average academic achievement of the sample, (e) the average socioeconomic status of the sample, (f) the percentage of minority students within the sample, (g) the primary language of participants, (h) the percentage of students with a disability, (i) the type of school in which the research was conducted (e.g., parochial, public), and (j) the country in which the evaluation took place.

Program features. A second series of items was developed to assess the variability of program features implemented across studies. Items included (a) the total length of the intervention (e.g., 8 weeks), (b) the average length of each session (e.g., one hour), (c) the frequency of intervention sessions (e.g., once a week), (d) the individual responsible for implementing the program (e.g., teacher, school psychologist), (e) the methods of training program implementers, (f)

the number of training workshops provided, and (g) whether or not booster sessions were provided to program implementers.

Methodological features. The final series of items were designed to examine the methodological features of eligible studies to determine areas of strength and weakness. Items within this domain included (a) the method of assignment (e.g., random, matching), (b) the unit of assignment (e.g., school, classroom), (c) the unit of statistical analysis (e.g., student, classroom), (d) whether or not the family-wise error rate was controlled for, (e) whether or not there was differential attrition from treatment or control groups, (f) whether or not social validity was reported, (g) whether or not treatment fidelity was reported, (h) the use of measures which could yield reliable and valid scores, (i) whether or not a follow-up assessment was conducted, (j) the use of multiple assessment methods, (k) the use of multiple informants to assess treatment effects, (l) the establishment of group equivalence at pretest on posttest measures, and (m) whether or not there were identifiable components linked to the primary outcomes.

Meta-Analytic Evaluation of the FRIENDS Program

A meta-analytic evaluation of the FRIENDS program was conducted for included studies to supplement the descriptive review of study, program, and methodological features. The purpose of this quantitative synthesis was to determine whether school-based applications of the FRIENDS program have demonstrated behaviorally-significant effects on standard measures of anxiety in student populations. In order to assist the research team in drawing conclusions regarding the efficacy of the FRIENDS program, effect size statistics were calculated to compare the posttest outcomes on standard measures of anxiety for students enrolled in treatment and control groups. These effect sizes were organized into two distinct analyses based on the risk level of sampled students. The first set of analyses was designed to evaluate the overall effectiveness of the FRIENDS program for preventing the onset of anxiety in students at low risk for developing anxiety disorders. Students were classified as low-risk if their pretest scores on standardized measures of anxiety were reported to be within the subclinical range. Effect sizes comparing students at low-risk were computed for immediate posttest and subsequent follow-up periods. A second series of analyses was undertaken to determine the effect of FRIENDS on students at high risk for developing anxiety disorders. Students were classified as high-risk if their pretest scores on standard anxiety measures were reported to be within the clinical range. As with the analyses focused on students at low-risk, effect sizes from posttest and follow-up periods were examined.

Organization of effect sizes. An important consideration in meta-analysis is to ensure that each distinct analysis contains an independent set of effect sizes (Lipsey & Wilson, 2001). As such, the following steps were taken to organize the various effect size statistics drawn from the studies included in the present review. First, many of the included studies reported results for two or more standard anxiety measures (see Table 1). Because preliminary sensitivity analyses comparing overall effect size estimates using randomly selected anxiety measures demonstrated statistically similar results, the research team randomly selected a representative anxiety measure for each study to include in final analyses. Second, some of the included studies compared results from two separate treatment groups that received different variations of the FRIENDS program against a single control group. Although the effect sizes for subsamples of the same study likely have shared dependencies (Landman & Dawes, 1982; Wolf, 1990), the convention in meta-analysis is to assume these dependencies to be trivial and to include them as independent samples (Lipsey &Wilson, 2001). Third, many of the effect sizes computed from follow-up studies were naturally based on repeated sampling of the same groups of students. For instance, the same group of participants might provide data at 12-, 24-, and 36-month post-intervention periods. In order to avoid the grouping of follow-up effect sizes drawn from the same sample of students within the same analyses, specific follow-up intervals were developed for low- and high-risk groups to assess the sustainability of treatment effects over time. Due to differences in the availability of follow-up data, comparisons for low-risk students included those drawn from periods within (a) the first 12 months and (b) beyond 12 months whereas high-risk follow-up analyses were only available for within 12 months following the completion of the FRIENDS curriculum.

Computation and correction of effect sizes. A series of standardized mean difference effect sizes were computed to describe the magnitude of treatment effect across treatment and control groups for the analyses previously described. All effect sizes for the present report were computed from the means and standard deviations reported within primary studies for relevant outcome measures. The sample sizes and variances associated with treatment and control groups were examined to ensure comparability. As such, the use of a standardized mean difference effect size with a pooled standard deviation remained viable. Initial effect size computations resulted in a Cohen's d which is defined as the difference between the mean outcome of the intervention group and the mean outcome of the comparison group divided by the pooled within-group standard deviation on the outcome mea-

Table 1
General Characteristics of Studies Included in the Present Review

Study	Sample Size	Participant Description	Country	Program Type	Effect Size Types*	Outcome Measures	School Type
S1 Barrett (2000)	n = 17	Mean Age = 16.30 Grades = Elem, HS % Male = 0	Australia	Selective prevention	Low-Risk Posttest	SCAS	Public
S2 Barrett (2001)	n = 204	Mean Age = 12.42 Grades = Elem, HS % Male = 52.45	Australia	Selective prevention	Low-Risk Posttest	RCMAS TSCL_Anx	Public
S3 Barrett (2003)	n = 320	Mean Age = 12.65 Grades = Elem, HS % Male = 52.19	Australia	Selective prevention	Low-Risk Posttest	RCMAS	N/R
S4 Barrett (2005)	n = 542 (posttest) n = 530 (12 month)	Mean Age = N/R Grades = 6, 9 % Male = N/R	Australia	Universal prevention	Low-Risk Posttest Low-Risk Follow-up Elevated Risk Posttest	SCAS	N/R
S5 Barrett (2006)	n = 669 (12-month) n = 430 (24 month) n = 226 (36 month)	Mean Age = 11.5 Grades = 6, 9 % Male = 48.60	Australia	Universal prevention	Low-Risk Follow-up	RCMAS SCAS	N/R
S6 Barrett & Turner (2001)	n = 588	Mean Age =10.75 Grades = 6 % Male = 50.51	Australia	Universal prevention	Low-Risk Posttest	RCMAS SCAS	N/R
S7 Bernstein (2005)	n = 61	Mean Age = 9.00 Grades = Elem % Male = 34.43	United States	Selective prevention	Low-Risk Posttest	MASC	N/R

Table 1 (continued)

Study	Sample Size	Participant Description	Country	Program Type	Effect Size Types*	Outcome Measures	School Type
S8 Hunt (2009)	n = 228 (24-month) n = 189 (48-month)	Mean Age = 12.05 Grades = 7 % Male =56.96	Australia	Selective prevention	Low-Risk Follow-up	RCMAS SCAS	Private
S9 Liddle (2010)	n = 58	Mean Age = N/R Grades = P5 – S3 % Male = 53.45%	Scotland	Selective prevention	Low-Risk Posttest	SCAS	N/R
S10 Lock (2003)	n = 737 (posttest) n = 636 (12-month)	Mean Age =N/R Grades = 6, 9 % Male = N/R	Australia	Universal prevention	Low-Risk Posttest Low-Risk Follow-up	RCMAS SCAS	N/R
S11 Lowry-Webster (2001)	n = 594	Mean Age = N/R Grades = 5-7 % Male = 47.14	Australia	Universal prevention	Low-Risk Posttest Low-Risk Follow-up Elevated-Risk Posttest	RCMAS SCAS	Private
S12 Lowry-Webster (2003)	n = 594	Mean Age = N/R Grades = 5-7 % Male = 47.14	Australia	Universal prevention	Low-Risk Follow-up Elevated-Risk Follow-up	RCMAS SCAS	Private
S13 Miller (2011a)1	n = 191	Mean Age = 10.1 Grades = 4-6 % Male = 52.00	Canada	Selective prevention	Elevated Risk Posttest	MASC	Public
S14 Miller (2011b)1	n = 253	Mean Age = 9.8 Grades = 4-6 % Male =46.00	Canada	Universal prevention	Low Risk Posttest	MASC	Public

Table 1 (continued)

Study	Sample Size	Participant Description	Country	Program Type	Effect Size Types*	Outcome Measures	School Type
S15 Mostert (2008)	n = 46	Mean Age = 12.50 Grades = 6 % Male = 52.00	South Africa	Universal prevention	Low-Risk Posttest Low-Risk Follow-up	SCAS	N/R
S16 Rose (2009)	n = 52	Mean Age = N/R Grades = 4 % Male = N/R	Canada	Universal prevention	Low-Risk Posttest	MASC	N/R
S17 Siu (2007)	n = 47	Mean Age = 8.4 Grades = 8-10 % Male = 52.17	Taiwan	Selective prevention	Elevated Risk Posttest	SCARED	Public

Note: MASC = Multidimensional Anxiety Scale for Children; RCMAS = Revised Children's Anxiety Scale; SCARED = Screen for Child Anxiety Related Disorders; SCAS = Spence Children's Anxiety Scale; TSCL_ANX = Trauma Symptom Checklist for Children – Anxiety Subscale. 1Miller (2011) contained two eligible studies that were coded separately for the present review.

sure (Borenstein, Hedges, Higgins, & Rothstein., 2009). This formula was not used for one of the included studies since the authors failed to ensure group equivalence through either random assignment of participants to treatment conditions or reporting on comparisons of initial group differences (Rose, Miller, & Martinez, 2009). As such, an alternative formula was used to derive *d* in which the initial group differences on pretest were taken into account (Lipsey & Wilson, 2001).

Following the computation of Cohen's *d* effect size statistics, a series of corrections were applied to increase their precision. For instance, study authors often assigned intact classrooms or schools to treatment or control conditions. When assignment is made at the aggregate level, the student-level sample size is not appropriate for calculating the variance and weighting function because it exceeds the number of independent observations. In such cases, the Cohen's *d* statistics and standard errors from these studies were adjusted following the procedures described by Hedges (2007). The adjustment requires specification of the proportion of variance shared by students nested within the same cluster, also known as the intraclass correlation (ICC). Unfortunately, the ICC is often not reported in studies using misaligned analyses. As such, the research team used a default ICC of .10 that is recommended by the What Works Clearinghouse (WWC; 2008) for behavioral and attitudinal measures for all corrections. It should be noted that the cluster adjustments have minimal effect on the effect size estimates themselves while attenuating the inverse variance weights used in the main effects and moderator analyses. These cluster-corrected effect sizes were further refined by applying Hedges' small sample size correction for all studies. As a result, the outcome measure for the present synthesis was Hedges' *g*.

The effect size *g* can range from zero to plus or minus any number of standard deviations depending on the direction and magnitude of the treatment effect. The conventions of effect sizes hold that an effect of .20 is considered small, a value of .50 is moderate, and values exceeding .80 are large (Lipsey & Wilson, 2001). For the present meta-analysis, effect sizes were expressed in such a way that negative effect sizes indicated reductions in anxious symptomatology. The distribution of effect sizes for each set of analyses was examined through stem and leaf plots to visually assess the normality of the data and to identify the presence of any outlying estimates of treatment effects. There were no outliers found for the present analyses.

Analysis of effect size statistics. The resulting effect sizes were combined and analyzed using a mixed effects framework in which it is assumed that the variance across treatment effects is derived from (a) subject-level sampling error, (b) systematic factors that can

be measured and modeled, and (c) random sources that were not, or cannot, be measured (Lipsey & Wilson, 2001). Subject-level sampling error refers to the natural residual variance associated with drawing a sample from a broader population. Systematic factors refer to those identifiable study features that might moderate the level of treatment effects such as subject characteristics or program variations. Random sources include a range of factors that are often beyond the scope of the study such as the resources available to students and the level of administrative support provided to teachers. In terms of the statistical implications of selecting a mixed-effects model, the variation not associated with subject-level sampling error is divided into two components, one representing systematic relationships between study characteristics and observed treatment effects and another component representing random study-level differences. This division of variability allows for systematic differences among studies to be modeled through moderator analyses while a random component of residual variation remains to represent statistical uncertainty.

An important issue in meta-analysis is not only to provide a summary of overall treatment effect but to also describe the consistency (or homogeneity) across observed effect sizes (Borenstein et al., 2009). The degree of homogeneity is an indication that differences among effect sizes might be related to factors other than subject-level sampling error and provides a context for interpreting the effects of study characteristics on treatment outcomes. The following methods were used to evaluate the degree of heterogeneity across overall effect size estimates with conclusions drawn from the consistency observed across these indices. These assessments included the Q-statistic which provides the ratio of observed to expected variation and, if significant, indicates the presence of heterogeneity, and the I^2 index which is interpreted as the proportion of variance reflected across observed effect sizes and is less sensitive to issues of the number of studies included in the analysis than other statistics.

Reliability

Reliability estimates were calculated for both the screening and coding of studies. In terms of study screening, 100% of potentially eligible articles were screened by both study authors. Each of the 49 eligible studies was categorized as either being included or not included in the review. The reliability for screening was 96.73%. Articles in which there was a discrepancy were reviewed by both authors and a consensus was derived. In terms of calculating the reliability for study coding, a total of 9 (52.94%) of the included articles were independently coded by study authors. Each portion of the coding manual was assessed on an item-by-item basis to derive an overall

reliability estimate for each of the three areas evaluated through the coding manual. Results of this reliability assessment revealed a high degree of agreement for items relating to (a) participant and setting characteristics (M = 97.68%; *Item Range* = 92.38% -100.00%), (b) program features (96.43%; *Item Range* = 94.58% - 100.00%), and (c) methodological features (98.56%; *Item Range* = 91.17% - 100.00%) of coded studies. In the case of discrepancies, articles were co-reviewed by the research team to identify the appropriate response.

Results and Discussion

The purpose of this review was to evaluate the effectiveness of school-based applications of the FRIENDS program for preventing the onset of anxiety in student populations to determine whether sufficient empirical evidence exists to support its use in schools and classrooms. Such research is important given that FRIENDS has been commercially marketed as the only evidence-based prevention program focused on anxiety and has garnered the endorsement of the WHO (2004). In order to conserve space, we present findings and discuss their implications for research and practice together in the following sections.

Sample and Program Characteristics

After screening of 49 potentially eligible studies, a total of 17 met inclusion criteria and served as the basis for the present meta-analysis. An overview of sample and program characteristics for these studies is presented in Table 1. A notable limitation regarding sample characteristics was the failure of study authors to provide participant information beyond age, gender, and school level. As such, there was a paucity of information known regarding potentially important moderating variables such as (a) the general academic achievement of sampled students, (b) the prevalence and types of disabilities of included students, and (c) the overall socioeconomic status of student participants. A summary of program features for each study is provided in Table 2. Significant trends relating to program features included (a) the wide range of roles associated with individuals charged with program implementation and (b) the reliance on workshops and ongoing technical support to assist with program implementation. These trends raise questions regarding the feasibility of implementing the FRIENDS program in school-based settings which are discussed in greater detail below.

Methodological Strengths and Weaknesses

The methodological features of studies included in the present review were assessed across 10 criteria. A summary of methodological

Table 2
Program Features of Studies Included in the Review

Study	Program Implementer	Training Type	Length of Initial Training	Length of Intervention	Intervention Session Intensity	Booster Sessions[a]
S1	Psychologist and Staff	Training Workshop Ongoing Support	3 days	10 weeks	N/R	N/R
S2	Mental Health Professionals	Training Workshop	N/R	N/R	60min / sess.	N/R
S3	Mental Health Professionals	Ongoing Support	N/R	N/R	60min / sess.	N/R
S4	Psychologists	Training Workshop Ongoing Support	N/R	10 weeks	45-60min / wk	Yes
S5	Teachers and Psychologists	Training Workshop	1 day	10 weeks	70min / wk	Yes
S6	Teachers and Psychologists	Training Workshop	1 day	10 weeks	75min / wk	Yes
S7	Psychologists	Training Workshop Ongoing Support	N/R	9 weeks	60min / wk	Yes
S8	School Counselor and Teacher	Training Workshop	2 days	10 weeks	50min / wk	Yes
S9	Research Team	N/R	N/R	10 weeks	N/R	N/R
S10	Psychologists	Training Workshop	1 day	10 weeks	75min / wk	Yes

Table 2 (continued)

Study	Program Implementer	Training Type	Length of Initial Training	Length of Intervention	Intervention Session Intensity	Booster Sessions[a]
S11	Teacher	Training Workshop Ongoing Support	2 days	10 weeks	60min / wk	Yes
S12	Teacher	Training Workshop Ongoing Support	1 day	10 weeks	75min / wk	Yes
S13	School Counselor and Teacher	Training Workshop Ongoing Support	1 day	9 weeks	60 min / wk	None
S14	School Counselor and Teacher	Training Workshop Ongoing Support	1 day	9 weeks	60 min / wk	None
S15	N/R	N/R	N/R	5 weeks	2 hour-long sessions / week	None
S16	Teacher	N/R	N/R	8 weeks	60min / wk	N/R
S17	Psychology Graduate Students	Training Workshop	N/R	8 weeks	N/R	None

Note: N/R = Not reported; [a]All booster sessions were reported to be a total of two sessions delivered at the first and third months following the ending of treatment.

features for each study is provided in Table 3. In general, the studies had some positive methodological features. For instance, all but one study either used random assignment or established group equivalence up front; treatment fidelity was reported in a majority of studies; and most studies used multiple methods to assess treatment success. However, there were also some consistent methodological weaknesses across studies including (a) the failure of research teams to control for family-wise error rates, (b) the lack of reporting on, and subsequent failure to control for, differential attrition between treatment and control groups for some studies, (c) the failure to report social validity in a majority of studies, and (d) the fact that most studies have been conducted by program developers. Three additional methodological issues that have implications for interpreting the subsequent meta-analytic results will be highlighted in greater detail.

The first, and perhaps greatest, methodological limitation of the studies included in the present review is the failure of all but three to report findings using the appropriate unit of analysis. By failing to appropriately account for dependencies among observations of students clustered within the same classroom, study authors overestimated the tests of statistical significance thereby increasing the likelihood of finding positive results (Thomas, Ramsay, & McAuley, 2003. While measures were taken in the current meta-analysis to correct for issues related to misaligned analyses, these procedures are sensitive to the value of the ICC used (Borenstein et al., 2009). As such, the following meta-analytic findings represent an approximation of overall treatment effectiveness with additional studies or reevaluations of previous work required to correct for mismatched analyses.

A second area of methodological concern relates to the extensive reliance on measures of student self-report to assess treatment outcomes. Although standard measures of anxiety were used for the present meta-analysis, the use of student self-report raises some questions regarding the reliability of research findings. For instance, research has demonstrated that scores from self-report measures of anxiety are often unable to reliably distinguish between students with and without anxiety (Campbell & Rapee, 1996; Viana, Rabian, & Biedel, 2008). The subsequent analyses on students at elevated risk, therefore, might be attenuated by the presence of some students at low risk for developing anxiety. Moreover, scores from self-report measures on anxiety have been shown to have poor test-retest reliabilities indicating instability over lengthy periods of time (Schniering, Hudson, & Rapee, 2000). As such, the data drawn from follow-up collection periods may not provide an accurate assessment of treatment effects.

A third area of methodological weakness related to the pervasive

use of waitlist control conditions in which participants in the comparison group received the treatment after some time delay while being provided business-as-usual care in the interim. Although the waitlist control is among the most prevalent comparison conditions used in behavioral research (Lindquist, Wyman, Talley, Findorff, & Gross, 2007), it is generally considered weak given its inability to control for nonspecific features of the intervention such as expectancy, social support, and attention (Vickers & de Craen, 2000). In terms of the FRIENDS evaluations reviewed here, the problems with the waitlist control were compounded by the lack of reporting on curricular and management protocols implemented in control classrooms. Moreover, the only two studies that used a control condition that accounted for generic features of the treatment found no effects of the program (Miller et al., 2011). These findings call into question whether observed treatment effects are, in fact, due to the proposed theoretical mechanisms suggested by program developers. Examples of designs that provide a more rigorous test of the active components of the intervention include randomized experiments with control students receiving either an alternative or placebo treatment.

Quantitative Evidence of the FRIENDS Program

The overall effectiveness of the FRIENDS program was assessed by calculating a series of standardized mean difference effect sizes comparing treatment and control groups on posttest and follow-up outcomes on measure of anxiety. Students were identified as either being at low or elevated risk based on their pretest scores on standard measures of anxiety. The results of these calculations are presented in Table 4 and are further described in the following sections.

Effects for low-risk students. The posttest outcomes for low-risk samples revealed that students enrolled in classrooms in which the FRIENDS program was implemented demonstrated a slight reduction in anxiety immediately following program completion. These small initial treatment effects were found to vary systematically (Q = 176.56, $p < .001$; $I^2 = 93.20\%$) although a series of moderator analyses of student characteristics and program features failed to predict treatment outcomes.

Secondary analyses for immediate posttest effects for low risk were subsequently conducted after removing those studies with poor methodological quality. Poor methodological quality was defined as those studies meeting fewer than half of the methodological criteria assessed as presented in Table 3 (i.e., a score of 4 or below). The findings of this secondary analysis resulted in the overall estimated level of treatment effect being significantly reduced (g = .12) and becoming

Table 3
Methodological Characteristics of Studies Included in the Review

	Study Number																	
Methodological Criterion	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	*Total*
Random Assignment	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	*12*
Appropriate Unit of Analysis	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	*3*
Type I Error Rate Controlled	1	0	1	1	1	-	1	-	1	1	0	0	1	1	-	-	0	*9*
Multiple Assessment Methods	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	*15*
Multiple Assessment Sources	0	0	0	0	0	0	1	0	1	0	1	1	1	1	0	0	1	*7*
Group Equivalence Established	1	1	1	0	0	1	1	1	1	1	0	0	1	1	1	0	0	*11*
Equivalent Attrition Reported	1	0	0	1	1	0	1	1	1	1	1	0	1	1	1	0	0	*11*
Social Validity Assessed	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	*4*
Treatment Fidelity Reported	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0	*13*
Independent Evaluation	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	8
Total	6	4	4	5	5	4	7	6	8	6	5	5	9	9	4	2	4	

Note: Criteria were developed from the coding manual developed by the American Psychological Association's Division 16 Task Force on Evidence-Based Interventions (Kratochwill et al, 2002.).

Table 4
Meta-Analysis Results for Studies Included in the Review for Anxiety Measures

Analysis	ES_n[a]	*g*	95% CI
Posttest Effect for Low-Risk Students	13	-.26*	-.48 – -.05
Follow-Up Effects for Low-Risk Students			
Up to 12 Months	6	-.31*	-.58 – -.05
Beyond 12 Months	2	-.09	-.19 - .02
Posttest Effect for Students with Elevated Risk	5	-.37	-.88 – .15
Follow-Up Effects for Students with Elevated Risk			
Up to 12 Months	5	-.21	-.49 – .07

Note: [a]The total number of effect sizes used to estimate overall treatment effect; [b]An outlying effect size estimate from Barrett and Turner (2001) was removed from this estimate. *Estimate was significant at the > .05 level; **Estimate was significant at the >.001 level.

non-significant (*CI* = -.45 - .19) indicating that some of the observed effects associated with the original analysis might have been due to alternative factors.

Regardless of whether methodological quality influenced the findings associated with immediate posttest results, low-risk students exposed to treatment were found to have lower levels of anxiety at follow-up data collection periods occurring within 12-months of program completion than students not exposed to treatment. These findings are associated with a significant degree of heterogeneity ($Q = 40.72$, $p = .001$; $I^2 = 90.20\%$), though moderator analyses again failed to explain any of this variance. There were an insufficient number of studies with low methodological quality for these analyses to warrant investigating the effects of study quality on overall levels of observed treatment effect. Finally, the effects of FRIENDS were not found to maintain for follow-ups occurring after 12 months, suggesting that the impact of the program on measures of anxiety deteriorated over time for low-risk students.

Effects for high-risk students. There were no differences observed for students with elevated risk enrolled in treatment and control classrooms on immediate posttest measures of anxiety. Moreover, these effect size estimates were found to be stable ($Q = 2.52$; $p = .28$; $I^2 = 20.70\%$), indicating that the absence of treatment effects was consistent across studies. Results drawn from follow-up data collected within 12 months of treatment completion revealed no differences between

groups for anxiety and were also consistent across studies ($Q = 5.56$, $p = .19$, $I^2 = 26.72\%$).

Implications for Research and Practice

The general conclusion of the present review is that the research underlying school-based versions of the FRIENDS program presently lacks the rigor to certify it as an evidence-based practice. These findings are in contrast to previous research generally suggesting that the program is effective (Horowitz & Garber, 2006; Schoenfeld & Janney, 2008; Stallard, 2010). The reasons for the disparate findings are likely related to the failure of previous reviews to meta-analytically combine treatment effects across studies and to consider the estimated magnitude in the context of methodological rigor.

Although results of the present review indicated that program exposure does seem to lead to initial reductions in anxiety for students at low risk for developing anxiety disorders, the methodological rigor of most studies included in the review undermines these general findings. As such, additional research on the FRIENDS program implemented as a school-based prevention program is needed to address the methodological weaknesses previously identified (Briesch, Hagermoser Sanetti, & Briesch, 2010). Despite the need for more rigorous research to validate the efficacy of school-based versions of FRIENDS, the commercial marketing and international recognition of the program as an evidence-based curriculum has led to its adoption in many school districts (Barrett & Pahl, 2006). As such, the research team encourages school personnel to carefully weigh the financial and temporal costs associated with implementing the FRIENDS program prior to adopting it to prevent the onset of anxiety in student populations. Given that many schools are presently implementing the program, however, we have identified three practical implications to enhance, potentially, the effectiveness of FRIENDS.

First, school personnel should develop a clear plan for providing support to students at elevated risk to cope with their anxiety. These support mechanisms can be manifested through methods of proactive screening and systematic, individualized interventions (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). In terms of screening, a recent push has been made to develop behavioral assessments with results that can reliably identify students with anxiety (e.g., Klein, Dougherty, & Olino, 2005; Schniering et al., 2000). These more focused measures might be used in conjunction with still-emerging screening procedures designed to assist school personnel in reliably identifying students who exhibit generalized symptoms such as social withdrawal and social skills deficits that are not solely

associated with anxiety (Cook et al., 2011). That is, brief, universal screening procedures might be used to identify candidates for more specific assessments in the interest of mitigating issues related to the under-identification of students with anxiety. Moreover, incorporating the prevention-focused FRIENDS program within a tiered intervention framework might provide a stronger context for students at the greatest risk to realize reductions in anxiety. Within a systematic approach to intervention, the prevention-focused FRIENDS program would ensure that all students receive some support for managing anxiety. More focused treatments that are provided in individual or small group formats would then be provided to students that continue to experience significant anxieties (Ollendick & King, 1998; Weisz, Hawley, & Doss, 2004). School personnel would subsequently be responsible for continually monitoring student responses to both preventive and selective treatments. By integrating the FRIENDS program within a systematic effort of screening, intervention, and progress monitoring, treatment effects for students at elevated risk may be realized sooner.

A second practical implication of the present review is the need for school personnel to support students at low-risk to maintain initial improvements in anxiety as compared to controls. According to this meta-analytic evaluation of the FRIENDS program, treatment effects for anxiety deteriorated at follow-up periods beyond 12 months. This abatement of treatment effects might be due, in part, to the inability or unwillingness of students to implement program procedures on their own as indicated by measures of social validity (e.g., Barrett, Sonderegger, & Sonderegger, 2001; Lowry-Webster, Barrett, & Lock, 2003). As such, those interested in using the FRIENDS program should consider adopting methods to continue student exposure to the curriculum or include formative skill assessments to ensure techniques are understood and used over time. Such methods might include (a) extending the curriculum past 10 weeks, (b) re-teaching critical program components via student-focused booster sessions, (c) developing methods for assisting students in generalizing the FRIENDS strategies to other contexts and situations, and (d) integrating those program mechanisms associated with the greatest levels of behavior change into school-wide initiatives. These extensions and augmentations of the FRIENDS program might assist with maintaining treatment results for students at low risk for developing anxiety disorders.

A third implication for practice is increasing the feasibility of implementing the FRIENDS program within classrooms. For instance, the predominant model used across reviewed studies was to implement the program with a school psychologist or counselor.

Although these school personnel might be the most appropriate individuals to put the FRIENDS program into practice given their training and expertise, many schools are constrained by time and finances to implement the program with integrity across several classrooms with such a model. The most efficient method for exposing students to FRIENDS, therefore, may be to train teachers to implement the curriculum. Interestingly, preliminary evidence has indicated that teachers (a) are able to implement the program with integrity, (b) find the program to be socially acceptable, and (c) may realize similar treatment gains for students as psychologists on measures of anxiety (Barrett & Turner, 2001; Lowry-Webster, Barrett, & Lock, 2003). Despite these promising results, additional research is needed to verify these findings and determine the feasibility of teacher implementation at the school level. Specifically, research is needed to identify those training methods and monitoring techniques that maximize teacher adherence to program components and subsequently lead to the greatest inhibition of student anxiety.

References

References marked with an asterisk indicate studies included in the meta-analysis. Please note that the 17 studies were drawn from 16 manuscripts since Miller et al. (2011) provided data from two eligible studies.

*Barrett, P. M., Farrell, L. J., Ollendick, T. H., & Dadds, M. (2006). Long-term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: An evaluation of the friends program. *Journal of Clinical Child and Adolescent Psychology, 35*, 403–411. doi:10.1207/s15374424jccp3503_5

*Barrett, P. M., Lock, S., & Farrell, L. J. (2005). Developmental differences in universal preventive intervention for child anxiety. *Clinical Child Psychology & Psychiatry, 10*, 539–555. doi:10.1177/1359104505056317

Barrett, P. M., Lowry-Webster, H., & Turner, C. (2000). *FRIENDS program for children: Group leaders manual.* Brisbane, STATE: Australian Academic Press.

Barrett, P. M., & May, S. (2007). *Introduction to FRIENDS: Anxiety prevention and treatment for children aged 7-11 and youth aged 12-16.* Retrieved from http://www.friendsinfo.net/ downloads/ FRIENDSintrobooklet.pdf

*Barrett, P. M., Moore, A. F., & Sonderegger, R. (2000). The FRIENDS Program for young former-Yugoslavian refugees in Australia:

A pilot study. *Behaviour Change, 17,* 124–133. doi:10.1375/bech.17.3.124.

Barrett, P. M., & Pahl, K. M. (2006). School-based interventions: Examining a universal approach to anxiety management. *Australian Journal of Guidance and Counseling, 16,* 55–75.

*Barrett, P., Sonderegger, R., & Sonderegger, N. L. (2001). Evaluation of an anxiety-prevention and positive-coping program (FRIENDS) for children and adolescents of non-English-speaking background. *Behaviour Change, 18,* 78–91. doi:10.1375/bech.18.2.78.

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

*Barrett, P. M., & Turner, C. (2001). Prevention of anxiety symptoms in primary school children: Preliminary results from a universal school-based trial. *British Journal of Clinical Psychology, 40,* 399–410. doi:10.1348/014466501163887

Beesdo, K., Knappe, S., & Pine, D. S. (2009). Anxiety and anxiety disorders in children and adolescents: Developmental issues and implications for DSM-V. *Psychiatric Clinics of North America, 32,* 483–524.

*Bernstein, G. A., Layne, A. E., Egan, E. A., & Tennison, D. M. (2005). School-based interventions for anxious children. *Journal of the American Academy of Child & Adolescent Psychiatry, 44,* 1118–1127. doi:10.1097/01.chi.0000177323.40005.a1

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* West Sussex, United Kingdom: John Wiley & Sons Ltd.

Briesch, A. M., Hagermoser Sanetti, L. M., & Briesch, J. M. (2010). Reducing the prevalence of anxiety in children and adolescents: An evaluation of the evidence base for the FRIENDS for Life program. *School Mental Health, 2,* 155–156.

Campbell, M. A., & Rapee, R. M. (1996) Current issues in the assessment of anxiety in children and adolescents: A developmental perspective. *Behaviour Change, 13,* 185–193.

Cartwright-Hatton, S., McNicol, K., & Doubleday, E. (2006). Anxiety in a neglected population: Prevalence of anxiety disorders in pre-adolescent children. *Clinical Psychology Review, 26,* 817–833.

Collins, K. A., Westra, H. A., Dozois, D. J. A., & Burns, D. D. (2004). Gaps in accessing treatment for anxiety and depression:

Challenges for the delivery of care. *Clinical Psychology Review, 24,* 583–616.

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine, 126,* 376–380.

Cook, C. R., Rasetshwane, K. B., Truelson, E., Grant, S., Dart, E. H., Collins, T. A., & Sprague, J. (2011). Development and validation of the student internalizing behavior screener: Examination of reliability, validity, and classification accuracy. *Assessment for Effective Intervention, 36,* 71–79.

Davis, T. E., May, A., & Whiting, S. E. (2011). Evidence-based treatment of anxiety and phobia in children and adolescents: Current status and effects of emotional response. *Clinical Psychology Review, 31,* 592–602.

Duchesne, S., Vitaro, F., Larose, S., & Tremblay, R. E. (2008). Trajectories of anxiety during elementary-school years and prediction of high school noncompletion. *Journal of Youth and Adolescence, 37,* 1134–1146.

Hedges, L. V. (2007). Correcting a significance test for clustering. *Journal of Educational and Behavioral Statistics, 32,* 151–179.

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

*Hunt, C., Andrews, G., Sakashita, C., Crino, R., & Erskine, A. (2009). Randomized controlled trial of an early intervention programme for adolescent anxiety disorders. *Australian & New Zealand Journal of Psychiatry, 43,* 300–304. doi:10.1080/00048670902721152.

Kauffman, J. M., & Landrum, T. J. (2009). *Characteristics of emotional and behavioral disorders* (9th ed.). Upper Saddle River, New Jersey: Pearson.

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

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merkangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62,* 593–602.

Klein, D. N., Dougherty, L. R., & Olino, T. M. (2005). Toward guidelines for evidence-based assessment of depression in children

and adolescents. *Journal of Clinical Child & Adolescent Psychology, 34,* 449–476.

Kratochwill, T. R., & Shernoff, E. (2004). Evidence-based practice: Promoting evidence-based interventions in school psychology. *School Psychology Review, 33,* 34–48.

Kratochwill, T. R., Stoiber, K. C., Christenson, S., Durlak, J., Levin, J. R., Talley, R., et al. (2002). In T. R. Kratochwill & K. C. Stoiber (Eds.), *Procedural and coding manual for review of evidence-based interventions.* Task force on the Evidence Based Interventions in School Psychology, Sponsored by Division 16 of the American Psychological Association and Society for the Study of Psychology.

Landman, J. T., & Dawes, R. M. (1982). Psychotherapy outcomes: Smith and Glass conclusions stand up under scrutiny. *American Psychologist, 37,* 504–516.

Layne, A. E., Bernat, D. H., Victor, A. M., & Bernstein, G. A. (2008). Generalized anxiety disorder in a nonclinical sample of children: Symptom presentation and predictors of impairment. *Journal of Anxiety Disorders, 23,* 283–289.

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45,* 163–191.

*Liddle, I., & Macmillan, S. (2010). Evaluating the FRIENDS programme in a Scottish setting. *Educational Psychology in Practice, 26,* 53-67. doi:10.1080/02667360903522785.

Lindquist, R., Wyman, J. F., Talley, K. M. C., Findroff, M. J., & Gross, C. R. (2007). Design of control-group conditions in clinical trials of behavioral intervention. *Journal of Nursing Scholarship, 39,* 214–221.

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

*Lock, S., & Barrett, P. M. (2003). A longitudinal study of developmental differences in universal preventive intervention for child anxiety. *Behaviour Change, 20,* 183-199. doi:10.1375/bech.20.4.183.29383

*Lowry-Webster, H., Barrett, P. M., & Dadds, M. R. (2001). A universal prevention trial of anxiety and depressive symptomatology in childhood: Preliminary data from an Australian study. *Behaviour Change, 18,* 36–50. doi:10.1375/bech.18.1.36

*Lowry-Webster, H. M., Barrett, P. M., & Lock, S. (2003). A universal prevention trial of anxiety symptomology during childhood:

Results at 1-year follow-up. *Behaviour Change, 20,* 25–43. doi:10.1375/bech.20.1.25.24843.

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

*Miller, L. D., Laye-Gindhu, A., Liu, Y., March, J. S., Thordarson, D. S., & Garland, E. J. (2011). Evaluation of a preventative intervention for child anxiety in two randomized attention-control trials. *Behaviour Research and Therapy, 49,* 1–9. doi:10.1016/j.brat.2011.02.00.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151,* 264–269.

*Mostert, J., & Loxton, H. (2008). Exploring the effectiveness of the FRIENDS program in reducing anxiety symptoms among South African children. *Behaviour Change, 25,* 85–96. doi:10.1375/bech.25.2.85.

Mychailyszyn, M. P., Mendez, J. L., & Kendall, P. C. (2010). School functioning in youth with and without anxiety disorders: Comparisons by diagnosis and comorbidity. *School Psychology Review, 39,* 106–121.

Neil, A. L., & Christensen, H. (2009). Efficacy and effectiveness of school-based prevention and early intervention programs for anxiety. *Clinical Psychology Review, 29,* 208–215.

Nemeroff, R., Levitt, J. M., Faul, L., Wonpat-Borja, A., Bufferd, S., Setterberg, S., & Jensen, P. S. (2008). Establishing ongoing, early identification programs for mental health problems in our schools: A feasibility study. *Journal of the American Academy of Child and Adolescent Psychiatry, 47,* 328–338.

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

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

Rubin, K. H., Burgess, K. B., Kennedy, A. E., & Stewart, S. L. (2003). Social withdrawal within childhood. In E. J. Mash & R. A. Barkley (Eds.). *Child psychopathology* (2nd ed., pp. 372–406). New York, NY: Guilford.

Rubin, K. H., Root, A. K., & Bowker, J. (2010). Parents, peers, and

social withdrawal in childhood: A relationship perspective. *New Directions for Child and Adolescent Development, 127,* 79–94.

Schoenfeld, N. A., & Janney, D. M. (2008). Identification and treatment of anxiety in students with emotional and behavioral disorders: A review of the literature. *Education and Treatment of Children, 31,* 583–610.

Schniering, C. A., Hudson, J. L., & Rapee, R. M. (2000). Issues in the diagnosis and assessment of anxiety disorders in children and adolescents. *Clinical Psychology Review, 20,* 453–478.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations and professional practices. *Journal of School Psychology, 45,* 193–223.

Shadish, W. R., Cook, T. D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Shadish, W. R., & Rindskopf, D. M. (2007). Methods for evidence-based practice: Quantitative synthesis of single-subject designs. *New Directions for Evaluation, 113,* 95–109.

*Siu, A. (2007). Using FRIENDS to combat internalizing problems among primary school children in Hong Kong: A pilot study. *Journal of Cognitive and Behavioral Psychotherapies, 7,* 11–26.

Stallard, P. (2010). Mental health prevention in UK classrooms: The FRIENDS anxiety prevention programme. *Emotional & Behavioural Difficulties, 15,* 23–35.

Thomas, R. E., Ramsay, C. R., & McAuley, L. (2003). Unit of analysis errors should be clarified in meta-analyses. *British Medical Journal, 15,* 393–397.

Tomb, M., & Hunter, L. (2004). Prevention of anxiety in children and adolescents in a school setting: The role of school-based practitioners. *Children & Schools, 26,* 87–101.

Tomblin, J. B., Zhang, X., Buckwalter, P., & Catts, H. (2000). The association of reading disability, behavioral disorders, and language impairment among second grade children. *Journal of Child Psychology and Psychiatry, 41,* 473–482.

Vaughn, M. G., Wexler, J., Beaver, K. M., Perron, B. E., Roberts, G., & Fu, Q. (2010). Psychiatric correlates of behavioral indicators of school disengagement in the United States. *Psychiatric Quarterly, 82,* 191–206. doi: 10.1007/s11126-010-9160-0

Viana, A. G., Rabian, B., & Beidel, D. C. (2008). Self-report measures in the study of comorbidity in children and adolescents with social phobia: Research and clinical utility. *Journal of Anxiety Disorders, 22,* 781–792.

Vickers, A. J., & de Craen, A. J. M. (2000). Why use placebos in clinical trials? A narrative review of the methodological literature. *Journal of Clinical Epidemiology, 53,* 157–161.

Weisz, J. R., Hawley, K. M., & Doss, A. J. (2004). Empirically tested psychotherapies for youth internalizing and externalizing problems and disorders. *Child Adolescent Psychiatric Clinics of North America, 13,* 729–815.

What Works Clearinghouse (2008). Procedures and standards handbook (Version 2.0). Washington, DC: Author. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf

Wolf, I. M. (1990). Methodological observations on bias. In K. W. Wachter & M. L. Straf (Eds.), *The future of meta-analysis* (pp. 139–151). New York, NY: Russell Sage Foundation.

World Health Organization (2004). *Prevention of mental disorders: Effective interventions and policy options.* Retrieved from http://www.who.int/entity/mental_health/evidence/en/prevention_of_mental_disorders_sr.pdf