PREVENTIVE MENTAL HEALTH AS AN APPROACH TO IMPROVING SCHOOL OUTCOMES AMONG YOUTH: A META-ANALYTIC REVIEW

Katrina Elaine Roundfield
DePaul University, kat.davis46@gmail.com

Follow this and additional works at: https://via.library.depaul.edu/csh_etd

Part of the Clinical Psychology Commons

Recommended Citation
Roundfield, Katrina Elaine, "PREVENTIVE MENTAL HEALTH AS AN APPROACH TO IMPROVING SCHOOL OUTCOMES AMONG YOUTH: A META-ANALYTIC REVIEW" (2014). College of Science and Health Theses and Dissertations. 80.
https://via.library.depaul.edu/csh_etd/80

This Dissertation is brought to you for free and open access by the College of Science and Health at Via Sapientiae. It has been accepted for inclusion in College of Science and Health Theses and Dissertations by an authorized administrator of Via Sapientiae. For more information, please contact digitalservices@depaul.edu.
PREVENTIVE MENTAL HEALTH AS AN APPROACH TO
IMPROVING SCHOOL OUTCOMES AMONG YOUTH:
A META-ANALYTIC REVIEW

A Dissertation Defense
Presented in
Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Clinical Psychology

BY
KATRINA ELAINE ROUNDFIELD
May 13\textsuperscript{th}, 2014

Department of Psychology
College of Science and Health
DePaul University
Chicago, Illinois
Dissertation Committee

Bernadette Sánchez, Ph.D.
Chairperson

Kathryn E. Grant, Ph.D.

Antonio J. Polo, Ph.D.

Eva Patrikakou, Ph.D.

William Sampson, Ph.D.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my dissertation chair, Dr. Bernadette Sánchez, for her support and guidance throughout my graduate career. I am so privileged to have experienced six years of Dr. Sánchez’s brilliant mentorship. To my dissertation committee, Drs. Kathryn Grant, Antonio Polo, Eva Patrikakou, and William Sampson, I am so honored to have each of you as members of my committee. Thank you for dedicating your time to reviewing my dissertation and sharing your expertise with me. I am sincerely indebted to my undergraduate research team and colleagues at DePaul, Drs. Sophia Duffy and David Meyerson, who helped me to complete this dissertation.

I am grateful to the Ford Foundation for supporting me as a pre-doctoral and dissertation fellow and to the Yale University School of Medicine Psychology Pre-doctoral Internship Program at The Consultation Center. My experiences as a Ford Fellow and at Yale have enriched my professional pursuits and have solidified my passion to advocate for children and other vulnerable populations.

Finally, I am truly blessed to have such wonderful family and friends who have been there every step of the way. My Mom, Dad, and brother, Kenny, have been my rock throughout my life. I could not have completed my dissertation without them. And to my husband, Stephen, you have made countless sacrifices to support my dreams, words cannot describe my gratitude.
VITA

Katrina (Davis) Roundfield was born in Oakland, California October 4, 1985. She graduated from Washington High School in 2003 and received her Bachelor of Arts degree from San Diego State University in 2008 and Masters of Arts degree from DePaul University in 2011. She completed her clinical psychology internship at Yale University School of Medicine in 2014.
Abstract

Researchers, policy makers, and educators continuously seek new avenues to enhance the academic achievement of children and adolescents. This goal is particularly pressing among youth from low-income, urban backgrounds, who are at increased risk for school failure (Aud, Wilkinson-Flicker, Kristapovich, Rathbun, Wang, et al., 2011). Taking a more holistic approach to understanding academic achievement, burgeoning research has begun to focus on the mental health of the child. Preventive mental health (PMH) is a theoretically sound and effective means of reducing the incidence of mental illness among youth from varying levels of risk (Durlak & Wells, 1997; 1998; Greenberg, Domitrovich, Bumbarger, 2000). The link between mental health outcomes and educational outcomes is well documented, such that reduced symptomatology is associated with better school outcomes for youth. The current investigation sought to meta-analyze PMH programs to determine whether these programs positively affect school outcomes among youth and to identify important study moderators.

The current investigation closely reviewed 142 studies yielded from search criteria and included a total of 35 studies with 46 independent samples in the meta-analysis. Results across 46 independent samples of universal, selected, and indicated PMH interventions yielded a small but significant effect size (Hedge’s g = .202). Proposed categorical and continuous moderators were not found to moderate program outcomes. The high variability in program approaches and measurement of school outcomes is implicated as a possible explanation for lack of significant moderators. This meta-analysis adds to the growing body of
literature that provides strong evidence for the causal relationship between mental health intervention and educational benefits for children and adolescents.
TABLE OF CONTENTS

Dissertation Committee ................................................................. ii
Acknowledgements ........................................................................ iii
Vita ................................................................................................ iv
Abstract ........................................................................................ v

CHAPTER I. INTRODUCTION ........................................................ 1
  Preventive Mental Health (PMH) .................................................... 2
  Mental Health and School Outcomes ........................................... 8
  The Role of Mental Health Interventions in School Outcomes ....... 17
  Moderators of Program Effects .................................................... 28
  Rationale for the Current Investigation ....................................... 34
  Research Questions and Hypotheses .......................................... 36

CHAPTER II. METHOD ................................................................. 39
  Inclusion/Exclusion Criteria ......................................................... 39
  Literature Search Procedures ...................................................... 42
  Coding Procedure ...................................................................... 44

CHAPTER III. DATA ANALYSIS .................................................... 47
  Computing Effect Size ................................................................ 47
  Analysis of Overall Program Effectiveness .................................. 48
  Moderator Analyses .................................................................. 50

CHAPTER IV. RESULTS ............................................................... 52
  Search Outcome ......................................................................... 52
  PMH Effects on School Outcomes ............................................. 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderators of Program Effects</td>
<td>69</td>
</tr>
<tr>
<td>Publication Bias</td>
<td>75</td>
</tr>
<tr>
<td>CHAPTER V. DISCUSSION</td>
<td>82</td>
</tr>
<tr>
<td>Mental Health, Mental Illness, and Positive Youth Development</td>
<td>83</td>
</tr>
<tr>
<td>Moderators</td>
<td>86</td>
</tr>
<tr>
<td>Limitations and Future Directions</td>
<td>94</td>
</tr>
<tr>
<td>Implications for Policy and Practice</td>
<td>97</td>
</tr>
<tr>
<td>References</td>
<td>100</td>
</tr>
<tr>
<td>Appendix A. Coding Guide</td>
<td>110</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1 Descriptive Characteristics of 34 PMH Interventions…………………53
Table 2 Study Characteristics…………………………………………………56
Table 3 Effect Sizes for Outcome Categories…………………………………68
Table 4 Results of Categorical Moderators…………………………………….72
Table 5 Results of Supplemental Categorical Moderators………………….75
LIST OF FIGURES

Figure 1 Effect Size Forest Plot Across 46 Independent Samples……..77
Figure 2 Funnel Plot of Standard Error by Standard Difference in Means
................................................................................................................78
Figure 3 Medialional Pathways from PMH to School Outcomes……..96
CHAPTER I

INTRODUCTION

Researchers, policy makers, and educators continuously seek new avenues to enhance the academic achievement of children and adolescents. This goal is particularly pressing among youth from low-income, urban backgrounds, who are at increased risk for school failure (Aud, Wilkinson-Flicker, Kristapovich, Rathbun, Wang, et al., 2011). These efforts have focused on increasing school funding, changing curricula, and altering elements of the school structure. Some of these efforts have reduced the gap in academic achievement among highly vulnerable groups (Becker & Luthar, 2002). However, the gap in achievement still remains (Aud, Wilkinson-Flicker, Kristapovich, Rathbun, Wang, et al., 2013; The Annie E. Casey Foundation, 2014).

Even beyond low-income, ethnic minority youth, educational goals in the U.S. still fall short. In 2011, the U.S. ranked 24th and 23rd out of 34 industrialized countries worldwide in math and science performance, respectively (Aud, Wilkinson-Flicker, Kristapovich, Rathbun, Wang, et al. 2013) and many of these countries are making educational gains in performance at twice the rate of the U.S. (Hanushek, Peterson, & Woessmann, 2012). Educational disparities between the U.S. and the international community have given rise to concerns that the U.S. will become less competitive in the global economy (Hanushek, Peterson, & Woessmann, 2012).

Taking a more holistic approach to understanding academic achievement, burgeoning research has begun to focus on the mental health of the
child. The report of the Surgeon General’s Conference on Children’s Mental Health (2000) formally acknowledged the critical link between children’s mental health and educational outcomes. In part because of this association, the report described the goal of increasing mental health functioning among all youth as a “national priority” (pp. 3). Becker and Luthar (2002) called for school reform to be focused on four critical components of academic achievement among disadvantaged youth specifically – school attachment, teacher support, peer values, and mental health. The three former components of academic achievement have been studied and incorporated into school reform efforts more systematically than the final component – mental health (Becker & Luthar, 2002). The continued exclusion of considering the role of mental health functioning in school outcomes demonstrates a gap in the educational reform efforts for our nation’s youth.

**Preventive Mental Health**

In the U.S., it has been estimated that 36.7% of youth will have been diagnosed with a mental or addictive disorder by the age of 16 (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003). Preventive mental health (PMH) programs for youth have long been established as an important means of reducing the incidence of mental illness (Weissberg, Kumpfer, & Seligman, 2003); Weisz, Sandler, Durlak, & Anton, 2005; Weisz, Weiss, Han, Granger, & Morton, 1995). Prevention is defined as an intervention that occurs before the onset of a disorder (Institute of Medicine, 1994). Within that definition exists several levels of prevention – universal, selective and indicative. The following sections will provide a theoretical framework for the role of PMH efforts in school outcomes
and explore strategies and empirical evidence related to PMH program effectiveness.

**Prevention Theory and Healthy Youth Development**

Prevention for youth problem behavior was developed during the 1970’s and 1980’s (Greenberg, Domitrovich, Bumbarger, 2001). Prevention programming emerged in response to difficulty in treating certain adolescent problem behaviors and the identification of specific antecedents of problem behaviors among youth (Ellis, 1998; Greenberg, Domitrovich, Bumbarger, 2001). These antecedents are called risk factors and must be considered developmentally – a risk factor for a 5 year old may differ from the risk factors present for a 17 year old. In addition, risk factors must be understood multi-systemically (Ellis, 1998; Greenberg, Domitrovich, Bumbarger, 2001). Guided by the ecological framework of Bronfenbrenner (1979), prevention research has emphasized the developmental importance of reciprocal interactions between the individual child and a person, object, or symbol within his/her environment. These interactions can occur at various levels of the ecological system – the individual, microsystemic, mesosystemic, or exosystemic levels (Bronfenbrenner, 1979).

Based on empirical evidence, Coie et al. (1993, p.1022), noted the following individual and environmental mental health risk factors among youth: constitutional handicaps (e.g., perinatal complications, neurochemical imbalance, organic handicaps, and sensory disabilities), skill development delays (e.g., low intelligence, social ineptitude, attention deficits, reading disabilities, and poor work skills and habits), emotional difficulties (e.g., apathy or emotional blunting,
emotional immaturity, low self-esteem, and poor emotional regulation), family circumstances (e.g., low social class, mental illness in the family, large family size, child abuse, stressful life events, family chaos and conflict, communication deviance, and poor parental bonding), interpersonal problems (e.g., peer rejection, alienation, and isolation), school problems (e.g., scholastic demoralization and school failure), and ecological risks (e.g., neighborhood disorganization, extreme poverty, racial injustice, and unemployment). The first three represent risk at the individual level, the following two involve the microsystemic level, and the final two represent mesosystemic and exosystemic risks. Notably, some of these risks are more malleable than others.

In contrast to risk factors, which are essentially probability markers associated with an increase in the likelihood of the onset of a particular problem or pathology, protective factors counteract the effects of risk (Dryfoos, 1990; Kia-Keating et al., 2011). Thus, protective factors serve as moderators, or buffers to disorder or dysfunction, or affect the mediational chain between risk and negative outcomes (Sandler, 2001). Also, protective factors may act directly to decrease dysfunction or to prevent the presence of a particular risk factor (Coie et al. 1993). Previous research identifying key protective factors for mental health comprise three domains: the child, the quality of the child’s interactions with the environment (microsystem), and aspects of the mesosystem and exosystem. First, at the individual child level, protective factors include cognitive skills, social-cognitive skills, temperamental characteristics, and social skills (Luthar & Zigler, 1992). Second, at the microsystemic level, interactions include secure attachments
to parents (Morissett, Barnard, Greenberg, Booth, & Speiker, 1990) and attachments to peers or other adults who engage in positive health behaviors and have prosocial values. Finally, at the mesosystemic and exosystemic levels, aspects such as school-home relations, quality schools, and regulatory activities may be protective for youth (Greenberg, Domitrovich, & Bumbarger, 2000).

**Universal, Selected, and Indicated PMH**

*Universal prevention* targeting mental health outcomes takes the approach that all youth may benefit from services in an effort to lower the incidence of mental disorders (Durlak & Wells, 1997). A more narrowly targeted universal approach, *selected prevention*, aims to select all youth who share a significant risk factor (Weisz, Sandler, Durlak, & Anton, 2005). An important consideration when distinguishing universal prevention from selected prevention approaches is to determine the selection criteria used to include program participants. Universal prevention does not discern which youths are at increased risk, while selected approaches target youths at elevated risk for developing a mental disorder. For instance, a universal approach may be to target all children involved in an after-school program. In selected prevention, however, the approach may be to target students experiencing a transition (e.g., transition to high school), youth who come from a single-parent household, or youth who have experienced violence in the home.

Strategies used in universal and selected PMH are often two-pronged: risk reduction and/or mental health promotion (Durlak & Wells, 1997). These prevention efforts often attempt to attain this goal by implementing programming
that provide youth with strategies and skills to reduce stress and risk behaviors and increase key competencies and coping skills (Weissberg, et al., 2003). Durlak and Wells (1997) conducted a seminal meta-analysis of universal and selected PMH studies involving youth ages 18 and under, published between 1970 and 1991. That meta-analysis focused on programs that were aimed at change within the individual child and programs aimed at environmental change (e.g., classroom management training for teachers, parent training). Findings from 177 outcome evaluations indicated that most universal prevention programs achieved significant positive effects (ESs ranging from .24 to .93). Youth were significantly more likely to have decreased mental health-related problems (e.g., anxiety, depression, behavior problems) and increased competencies (e.g., assertiveness, communication skills, feelings of self-confidence) compared to controls. These findings are particularly notable because these samples were universal and selective, suggesting that most of the study participants would be functioning normally (i.e., not clinical or sub-clinical samples), and significant changes in functioning would not have been entirely expected.

Another prevention approach is indicated prevention which targets a specific subset of the population who evidence greater risk of developing a mental disorder. The goal of this level of prevention is to intervene before full criteria for a clinical disorder manifests (Durlak & Wells, 1998). Unlike selected prevention, a systematic screening of a target population is typically conducted in indicated preventive intervention, and those who meet criteria (e.g., symptoms of depression but do not meet full criteria for major depressive disorder) are selected
from the population to be included in the intervention. These indicated samples are often referred to in the literature as “at-risk” of developing some mental disorder or psychosocial problem.

Strategies used in indicated PMH typically involve the adaptations of standard clinical treatments for particular disorders. Often times, these treatments aim to reduce risks and increase competencies, similar to universal and selective approaches. Durlak and Wells (1998) conducted another meta-analysis of 130 indicated PMH program studies between 1960 and 1991 among children under the age of 19 years. Indicative interventions that were behavioral (ES = .51) or cognitive-behavioral (ES = .80) in their prevention approach yielded large effect sizes. Non-behavioral indicative prevention programs were less effective (ES = .09). Authors concluded that indicated PMH is an effective means of reducing youth mental health problems and increase key competencies (Durlak & Wells, 1998).

**Summary: PMH**

PMH is a theoretically sound and effective means of reducing the incidence of mental illness among youth from varying levels of risk (Durlak & Wells, 1997; 1998; Greenberg, Domitrovich, Bumbarger, 2000). PMH seeks to reduce key developmental risks and increase developmental protective factors within a child and their contexts.

Due to the complexity of the multi-systemic developmental risk and protective factors, targeted PMH efforts may yield benefits that extend beyond the mental health arena (Greenberg, Domitrovich, & Bumbarger, 2000). Indeed, there
are multiple pathways to the development of mental illness, which makes it all the more imperative to target multiple risks and protective factors across multiple systems (Ellis, 1998; Greenberg, Domitrovich, & Bumbarger, 2000).

Additionally, many developmental risk factors are not disorder-specific, or even mental health specific, but instead are related to a range of maladaptive outcomes (Greenberg, Domitrovich, & Bumbarger, 2000). For instance, a child may experience a risk, such as low parental involvement, which is both a risk for mental illness (Coie et al., 1993) and academic failure (Fan & Chen, 2001). After a comprehensive review of the literature, Greenberg and colleagues (2000) suggested due to the overlap between risk factors and multiple outcomes (e.g., mental illness, social and school problems), PMH efforts that focus on reducing risks and promoting protective factors may have direct effects on other outcomes outside of mental health (Coie et al., 1993; Dryfoos, 1990). The following section will explore in more detail the relationship between mental health and school outcomes. Just as school problems is a risk factor for the development of mental illness (Coie et al., 1993; Dryfoos, 1990; Ellis, 1998), mental illness is a risk factor for the development of school problems (Resnick, 2000).

**Mental Health and School Outcomes**

Previous research has revealed that mental health is related to school outcomes (DeSocio & Hootman, 2004; Puskar, Sereika, & Haller, 2003; Zychinski & Polo, 2011). For the purposes of the current investigation, the term *school outcomes* will be used to describe a variety of outcomes (e.g., performance, classroom behavior, attendance, academic attitudes, etc.) that are
relevant to the academic success of youth. Youth who struggle with mental illness are more likely to drop out of school, earn poorer grades, and engage in delinquent behavior in and outside of the classroom compared to youth without mental illness (Fergusson & Woodard, 2002). The reverse is also true, such that youth who struggle in school are more likely to also exhibit externalizing behaviors or suffer from an emotional disorder (Puskar & Bernardo, 2007).

Given the prevalence of mental disorders among youth and the important link between mental illness and school outcomes, there is a growing need to bridge these two domains in order to provide quality care to youth. Innovative prevention approaches that help to prevent mental health problems while promoting school outcomes are warranted (Zychinski & Polo, 2011). However, prevailing mental health and education intervention research and practice rarely consider these areas as being related beyond a correlational nature. Little is known about whether prevention efforts targeting mental health outcomes can also be effective in promoting school outcomes among youth. In order to advance mental health preventive initiatives as well as educational policy it is critical to determine whether PMH efforts are able to cause lasting effects on educational outcomes.

**Patterns in Mental Health and School Outcomes**

Mounting evidence suggests the interdependence of school functioning and mental health (Masten & Coatsworth, 1998; Roeser, Eccles, & Strobel, 1998; Roeser, Eccles, & Freedman-Doan, 1999; Zychinski & Polo, 2011). A comprehensive 10-year, longitudinal investigation of the relationship between
psychological and school functioning was conducted (Roeser, Eccles, & Freedman-Doan, 1999). Participants (N = 184) were followed from elementary school (mean age = 7.32 years) through high school (mean age = 14.2 years). Youth self-reported on measures of achievement motivation, academic competence, academic values, and mental health (Symptom Checklist – 90). Additionally, teachers were asked to report youths’ academic and social competence, school records were gathered to assess children’s grades, and research assistants tested cognitive abilities by using the Slosson Intelligence Test (Slosson, 1963). Each of these measures was collected with the purpose of elucidating the relationship between mental health and school outcomes across childhood and adolescence, and particularly during the transition to high school.

Cluster analyses revealed four distinct groups of students. The first cluster represented youth who were well-adjusted (n = 47). These youth were characterized by high levels of school motivation and mental health. Well-adjusted youth earned significantly higher grades in comparison to the other groups. The second cluster, those with poor academic motivation only (n = 48), included youth who devalued school and felt academically incompetent despite reporting little emotional distress. These students earned average grades in school. The third group represented those with poor mental health only (n = 46). These students reported significantly higher academic achievement than the sample, but had poor mental health and high emotional distress. Finally, there were those who had multiple problems, representing nearly 25% of the sample (n = 43). These youth reported both low academic motivation and poor mental health. Students
with multiple problems earned significantly lower grades compared to the other clusters. All clusters were found to be relatively stable across the transition from middle school to high school. Even when controlling for cognitive ability in Grade 1, the clusters remained stable, suggesting that cognitive ability was not primarily responsible for discrepancies in academic performance across clusters. These analyses generally suggest that academic achievement and mental health problems are related, and these associations are stable over time.

This study was one of the most comprehensive longitudinal studies examining the relationship between mental health and academic achievement among youth. Authors emphasized the persistence of school and mental health problems even when controlling for previous academic functioning, suggesting that multiple problem youth in particular may be experiencing a dynamic in which strained mental health exacerbates poor academic outcomes, and vice versa (Roeser, Eccles, Freedman-Doan, 1999). Authors concluded that PMH may be one approach to reduce the negative cycle of poor mental health and poor academic performance through high school, as high school typically exacerbates existing problems (Dryfoos, 1990; Eccles et al., 1997). Roeser and colleagues’ (1999) study is just one example of how students in a general school setting may be struggling with mental health concerns and how these mental health issues may have a negative effect on academic outcomes.

Research focusing on youth who exhibit emotional or behavioral distress provide even further evidence of the link between psychological functioning and academics. Youth who may be struggling with general distress or sub-clinical
emotional disorders are more likely to experience difficulties in the classroom as a result. Some of these difficulties include social withdrawal, often expressed as difficulty participating in the classroom, trouble concentrating in class or on assignments, and the inability to engage in cognitive, psychomotor, and affective learning tasks (Puskar & Bernardo, 2006). For instance, using causal modeling techniques, Masten and colleagues (1998) examined cross-time relations of children’s school, social, and behavioral functioning through adolescence. Results indicated that in adolescence, conduct problems lead directly to difficulties in academic attainment.

Among youth who exhibit maladjustment consistent with sub-clinical levels of internalizing disorders, such as depression or anxiety, studies have shown these problems may affect school outcomes. For instance, there is evidence that depressed mood and anxiety may affect participation in school, due to a child’s depressed mood affecting their self-esteem (Puskar & Bernardo, 2006). Shedding light on the mediating factors associated with academic achievement and mental illness among youth, Zychinski and Polo (2011) found that academic self-efficacy mediated the relationship between depression and academic performance. Specifically, when examined as a mediator, self-efficacy resulted in a 29.6% reduction in the association between GPA and depression. Zychinski and Polo (2011) noted in their limitations that they did not present a causal model in their research design, lending two possible explanations. The first explanation is that low academic achievement may perpetuate depressive symptoms by decreasing youths’ academic self-efficacy. The other hypothesis is that
depression, via a less hopeful sense of oneself, may negatively affect academic self-efficacy, thus producing poorer academic performance outcomes. An implication of Zychinski and Polo’s (2011) study and studies that report similar bi-directional associations between mental health and school outcomes (e.g., Roeser, Eccles, Freedman-Doan, 1999) is that there is even greater benefit of prevention efforts to promote mental health among youth aside from just psychological benefits. That is, one may increase school functioning through enhancing mental health functioning, and, in turn, yield on-going mental health benefits associated with strong school functioning. However, in order to produce these psychological and school-related benefits, one must have a firm conceptualization of why the two areas are related.

Examining mental illness and mental health as a dual-factor model provides a better understanding of the relationship between mental health and school outcomes. Suldo and Shaffer (2008) examined a dual-factor model of mental illness and mental health and its relation to psychosocial functioning among adolescents. In the dual-factor model (Keyes, 2005), mental illness is defined as the presence of psychopathology – internalizing (e.g., depression, anxiety) and externalizing (e.g., conduct disorder, oppositional defiant disorder) disorders, while mental health is defined by subjective well-being. Subjective well-being was measured in this study by life satisfaction, and positive and negative affect (Diener, 2000). The overarching goal of this study was to explore the dual factor model among youth and the psychosocial outcomes associated with mental illness or health.
Participants were 349 students from grades 6 through 8, ranging from 10 to 16 years (M = 12.96). Youths’ school functioning (GPA, attendance, standardized test scores, and school attitudes), mental illness (i.e., Youth Self Report [YSR] internalizing symptoms, Teacher Report Form [TFR] externalizing symptoms), and mental health (i.e., positive and negative affect and life satisfaction) were assessed. Four distinct groups emerged from the data: (a) youth with complete mental health (57%) or both subjective well-being and the absence of mental illness; (b) vulnerable youth, those without psychopathology but low subjective well-being (13%), (c) symptomatic but content youth (13%), or those who exhibited high psychopathology scores yet reported average to high subjective well-being; and (d) troubled youth (17%) who had high psychopathology and low subjective well-being. Based on previous research (i.e., Keyes, 2005), students who were reported with low subjective well-being were identified as “languishing,” while those with high subjective well-being were identified as “flourishing” (Suldo & Shaffer, 2008). Between-subject Multivariate Analysis of Covariance (MANCOVA) tests were conducted to determine differences in school functioning across the four groups. As expected, school functioning was highest among the complete mental health group of students. Symptomatic but content youth were the second highest academically achieving subgroup. Interestingly, vulnerable youth (languishing only) and troubled (languishing and high psychopathology) youth were significantly lower than average in academic performance and did not significantly differ from one another in academic performance. Authors concluded that subjective well-being
helped to buffer the full negative effects of psychopathology on school outcomes (Suldo & Shaffer, 2008).

Several important implications emerged from Suldo and Shaffer’s (2008) study. Nearly half of the entire sample of students in their study was not at optimal psychological health, based on languishing, mental illness, or a combination of the two. Adolescents who were diagnosed as anything less than complete mental health (i.e., high subjective well-being without mental illness) were functioning worse in a number of domains, including school (Suldo & Shaffer, 2008). As further evidence for the importance of PMH, languishing-only adolescents function as poorly on most school outcomes as those with a mental illness. Suldo and Shaffer (2008) pointed to suggestions for PMH from their research findings, stating that traditional mental health programming would exclude these adolescents who are languishing because they are not diagnosable. Authors recommended undergoing screening procedures to capture those youth who are languishing but may not be exhibiting mental illness. However, because these developments in mental health and illness are relatively new, the current study will explore universal, selective, and indicative prevention efforts, assuming that languishing children and youth would be captured by more traditional preventive programming, rather than slipping undetected through the cracks of traditional mental health treatment interventions. Authors asserted that “fostering well-being in all children may be essential to attain maximum positive academic function” (p. 64). And “If mental health professionals and educators continue to
focus exclusively on identifying and treating mental illness, optimal functioning may be less likely for the majority of students.” (p. 66).

This expansion from simply understanding mental health as the absence of a mental illness allows for a clearer understanding of why mental health may help to facilitate a number of positive life outcomes. Suldo and Shaffer (2008) provided strong evidence for why optimal psychological functioning is important for all students as well as for those who are already mentally ill or at risk for mental illness. Thus, universal and indicated PMH efforts are of utmost importance to enhance school outcomes among youth.

**PMH & School Outcomes: Examining the Evidence**

Revisiting PMH theory provides more conceptual evidence that this approach may improve school outcomes. Resnick (2000) suggested threats to well-being, such as poor mental health or languishing, may have deleterious effects on youths’ developmentally appropriate functioning. Resnick defined developmental functioning as achievement of the tasks that are appropriate to that developmental age range. For children and adolescents, one of the most salient domains of developmental functioning is school (Resnick, 2000). Thus, academic achievement and school outcomes that are mediators of academic achievement (e.g., academic self-efficacy, school engagement) may be indicators of healthy development for many youth. To illustrate, if healthy development is compromised by poor mental health, such developmental indicators as academic achievement may also be negatively affected. Many of the studies mentioned above (e.g., Suldo & Shaffer, 2008; Roeser, Eccles, & Freedman-Doan, 1998;
Zychinski & Polo, 2011) provide evidence of this pattern of reciprocal
dysfunction between mental health and school outcomes. Again, this pattern is
more intuitive when mental illness and poor school functioning are understood as
risk factors of one another.

The Role of Mental Health Interventions in School Outcomes

The current investigation is a meta-analysis of PMH studies that report
school outcomes. The primary goal of this meta-analysis will be to synthesize and
analyze quantitative data from evaluations of PMH programs in order to
determine their effectiveness in increasing school outcomes among youth. No
meta-analyses to date have definitively sought to answer the question of whether
PMH produces positive school outcomes among youth. However, some studies
exist that had similar goals. The following section reviews previous studies that
have synthesized evaluation results of mental health interventions reporting
effects on school outcomes. In addition to reviewing these studies, this section
will highlight limitations of these works and provide a clear rationale for the
current investigation.

Previous Narrative Syntheses

A narrative synthesis, unlike a meta-analysis, does not test quantitative
data in order to draw over-arching conclusions (Cooper, 2010). Instead, narrative
synthesis uses the methodology of summarizing study characteristics in narrative
form and often provides conclusions based on counting significance of treatment
groups versus controls, or counting p values (Cooper, 2010). This methodological
concern presents two major limitations. First, results based on the counting of p
values should be considered with extreme caution. Numerous methodological reviews strongly suggest that reports of p values from individual studies may often be misleading, and that effect size (ES) values and meta-analyses that synthesize those values across multiple studies is more reliable (Cohen, 1990; Cooper, 2010; Lipsey & Wilson, 2002). In fact, a recent meta-analysis (Farahmand, Grant, Polo, Duffy, & Dubois, 2011) examining the effectiveness of school-based mental health programs among urban, ethnic minority youth provided direct contradiction to another narrative synthesis (i.e., Rones and Hoagwood, 2000) that used the methodology of counting p values. Through quantitative meta-analysis, Farahmand and colleagues (2011) found, contrary to Rones and Hoagwood’s (2000) conclusion, that school-based mental health was significantly less effective for low-income urban youth. The second methodological concern about narrative synthesis is that moderation analyses cannot be conducted because because there are no effect sizes to represent quantitatively the size of the change from pre- to post-test. That said, results from narrative reviews serve as important building blocks for generating hypotheses; two will be discussed in this section.

The first synthesis to date that explicitly sought to answer the question of whether mental health programs are effective in increasing school outcomes by exploring past studies was conducted by Hoagwood, Olin, Kerker, Kratochwill, Crowe and colleagues (2007). Hoagwood et al. (2007) conducted a narrative synthesis of school-based mental health prevention and treatment studies published between 1990 to 2006. Researchers reported that 24 (37.5%) studies
tested the effects of a mental health program on both school and mental health outcomes. In their narrative synthesis, Hoagwood and colleagues (2007) reported that both mental health and school outcomes were positively affected by school-based PMH programs. The most effective were multi-level intervention approaches in schools, those that included intervention beyond the child level (e.g., parent training), especially during transitional periods.

Hoagwood and colleagues’ (2007) review, yet informative and important, has several limitations. First, school outcomes were defined by behavioral markers only (i.e., grade point average [GPA], attendance, misconduct, and special education placement). This approach to measuring school outcomes is problematic in intervention research because behavioral change in the form of performance is unlikely to change in such short follow-up assessments (most evaluations reported immediate post-test results to 1–year follow-up). Other school-related measures (e.g., academic self-efficacy, school engagement) would provide a more rich picture of the changes as well as the mechanisms associated with PMH intervention and academic achievement. Second, Hoagwood and colleagues (2007) limited their review to just school-based mental health interventions. Although the majority of mental health interventions that document school outcomes occur in the school setting, it is unknown whether preventive programs that occur outside of the school setting may also yield positive benefits. Another important limitation of Hoagwood et al.’s (2007) study is related to the general limitations of a narrative synthesis compared to a meta-analysis. Formal moderation analyses could not be conducted. This means that conclusions about
the effectiveness of multi-systemic programs over individual-level programs, for example, cannot be fully trusted. The absence of moderation analyses also limits the full interpretation of findings (Cooper, 2010). For example, moderation analyses allowed Farahmand and colleagues (2011) to determine that low-income urban youth did not participate in treatment programs that held a high standard of fidelity. Farahmand et al.’s (2011) findings provided further evidence to question the trustworthiness of Hoagwood et al.’s (2007) conclusions about the educational benefits of school-based PMH for all youth, as well as the factors associated with effective programming.

The second narrative synthesis that focused on school outcomes associated with mental health intervention was conducted by Becker, Brandt, Stephan, and Chorpita (2014). In their review of mental health treatment (indicated prevention and full treatment of a DSM-IV disorder) interventions, they used a vote-counting method similar to Hoagwood and colleagues (2007). Becker et al. (2014) found that 83.3% of the 602 treatment groups reported positive school outcomes compared to comparison groups. These findings suggest that mental health treatments likely yield benefits outside of the mental health arena.

Becker and colleagues (2014) included a wide range of school outcome indicators and examined treatments in a number of settings (i.e., not just school-based). However, the current meta-analysis differs in a number of ways aside from the meta-analytic approach versus a narrative synthesis approach. Most notably, the current meta-analysis focused on prevention programs and excluded treatment studies. The current investigation adds to the systematic review
literature of mental health intervention effects on school outcomes by its meta-analytic approach and examination of moderators of study outcome effects.

**Previous Meta-Analyses**

Among meta-analyses, none have provided in-depth analyses of whether universal and indicated PMH programs are effective in promoting school outcomes. Several meta-analyses exist that have reported effect sizes of school outcomes after mental health intervention (e.g., Durlak & Wells, 1997; 1998; Prout & Prout, 1998). These investigations have reported effect sizes from 0.0 (Prout & Prout, 1998) to small (Durlak & Wells, 1997). Because these meta-analyses were not specifically focused on exploring school outcomes, the types of school outcomes were limited to performance outcomes only and did not focus moderation analyses on school outcome effect sizes. Additionally, these meta-analyses ranged from prevention to treatment interventions.

**Baskin and colleagues (2010)**. The only existing meta-analysis that is similar to the current investigation’s aims is a meta-analysis of psychotherapy interventions conducted by Baskin, Slaten, Sorenson, Glover-Russell, and Merson (2010). This study reported that treatment interventions targeting children and adolescents diagnosed with a mental health disorder were effective at increasing positive school outcomes (Baskin, Slaten, Sorenson, Glover-Russell, & Merson, 2010). The meta-analysis included 83 studies with 102 treatment comparisons conducted between 1980 to 2008 that reported both mental health and school outcomes. Psychotherapy interventions aimed to reduce mental health problems among youth or aimed to increase adaptive behaviors through counseling. PMH
interventions were excluded from the meta-analysis. School outcome measures were defined as instruments that assessed a construct directly related to the current or future academic performance of a youth (e.g., teacher-rated on-task behavior in classrooms, specific tests of academic ability, student-reported academic self-efficacy, grades, teacher rating of being a good student, self-rated attitude toward schoolwork, school attendance). The aggregated school outcome effect size was small (d = .38) but significant. Because the focus of this study was on the impact of psychotherapy treatment interventions on school outcomes, further analyses were conducted in order to categorize the variety of school outcome measures into more homogenous groups. Authors categorized school outcomes into teacher-rated classroom behavior (d = .26), direct assessment of academic performance (d = .36), environmentally related outcomes, such as attendance and discipline (d = 0.26), and self-reported school-related outcomes (d = 0.59). All school categories reached homogeneity and were significantly different from 0. Small to moderate school outcomes were explained as being expected, as there are many factors that affect school outcomes aside from mental health functioning. However, authors remained that psychotherapy may serve as a means of supporting youths’ school functioning (Baskin et al., 2010).

Moderation analyses revealed no differences in school outcome effect sizes based on gender. Developmental age did not significantly affect outcomes (adolescents reported stronger school outcomes d = .45; children d = .41 and studies coded as being mixed with both adolescents and children d = .32). Baskin and colleagues’ (2010) meta-analysis provided evidence for the efficacy of
psychotherapy treatment interventions among youth. Results further solidify the important role of mental health functioning in school success.

Several limitations are important to note about Baskin and colleagues’ (2011) analysis. Because this study was limited to psychotherapy interventions with children and adolescents already diagnosed, it is unknown whether PMH programs would yield similar results. Further, moderation analyses from this meta-analysis were very limited. Specification of the types of intervention characteristics that were most helpful in promoting school outcomes was not examined. Further, multi-component interventions were included in the meta-analysis but no analysis was conducted to determine whether multiple component programs outcomes differed from single component program outcomes. Previous studies have reported that multiple components may be particularly helpful in school-based mental health interventions with youth (e.g., Hoagwood et al., 2007). Finally, studies included in the meta-analysis varied across different treatment settings, such as schools, clinics, and home-based treatments. Analyzing whether settings affected school outcomes would be an important question to answer.

**Social-emotional learning meta-analyses.** Finally, two other meta-analyses provided evidence of the positive effects of mental health programming on academic outcomes. More specifically, these two meta-analyses focused on studies that evaluated the effectiveness of social-emotional learning interventions (SEL). SEL seeks to promote the following competencies in an effort to promote healthy psychological development of youth: self-regulation, self- and social
awareness, and decision-making and relationship skills (Greenberg et al., 2003). Meta-analyses that evaluated SEL prevention efforts included interventions that aimed to enhance at least one SEL competency. The first meta-analytic study evaluated over 213 school-based SEL interventions (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). All SEL programs included in the study were universal prevention programs, and the authors excluded any studies that sought to treat symptomatic or sub-clinical youth. Results indicated that SEL, when implemented properly (i.e., sequenced, active, focused and explicit [SAFE]), was mildly effective (d = .28) at producing positive academic performance (i.e., standardized tests, GPA, and grades) among youth ages 5 to 18 years. Moderation analyses revealed that multi-component programs, those that involved multiple levels of intervention beyond the individual child level, were not found to be superior to single-component programs. However, authors discussed that adhering to SAFE procedures was less likely in multi-component programs, thus providing a confounding variable in the effectiveness of multicomponent SEL programs.

The second meta-analysis of SEL programs focused exclusively on indicated preventive SEL programs and reported greater effect sizes for school-based programming compared to the meta-analysis of universal SEL programming (Payton, 2008). A total of 324 school-based and after-school setting indicated prevention studies conducted between 1970 to 2007 were included. In Payton’s (2008) meta-analysis of indicated prevention SEL programs targeting children ages 5 to 13 years, school-based SEL programs were found to have a
moderate (d = .43) and significant effect on academic performance. Programs that were implemented after-school in the community yielded a lower effect size (d = .08), yet this effect was statistically significant from 0.

Although these studies provide a great deal of evidence for the value of school-based mental health prevention programming, several gaps exist. One key point when considering meta-analyses of SEL programs is that, although similar to general PMH programs, there are conceptual discrepancies. SEL is essentially one subset of mental health promotion. SEL seeks to enhance skills, but does not necessarily include programs that seek to reduce symptoms specifically. This conceptual discrepancy yields both gaps and areas of conflation in the types of programs that were included in meta-analyses exclusive to SEL. For example, Durlak and colleagues’ (2011) meta-analysis included some cognitive-behavioral therapy (CBT) prevention efforts because CBT seeks to promote self-regulation strategies; studies that would likely be included in the current investigation. However, the study also included programs that seek to teach leadership skills to youth, which represent studies that would not be included in the current investigation. Second, similar to a Hoagwood et al. (2007) limitation, evaluations of universal SEL prevention programs have only been school-based. Thus, little is known about the effectiveness of universal SEL components outside of the school setting. Third, Durlak et al. (2011) nor Payton (2008) examined in-depth the various components involved in these prevention interventions, such as teachers, parents, or other important agents within the students’ environment. Thus, it is unclear whether constructs outlined in SEL programming are also effective
outside of the school setting or are more effective when coupled with reinforcing agents, such as teaching parents to encourage SEL skills at home. Finally, only a limited analysis of the role of SEL on academic outcomes was included. The examination of academic outcomes was a tertiary goal of the study thus this outcome was not examined in any depth (e.g., moderation analyses exclusive to academic effect sizes). Also, similar to Hoagwood and colleagues (2007), academic indicators were only performance related. Authors pointed out this limitation in their discussion and suggested that future studies examine more proximal outcomes (e.g., school attitudes) related to the academic realm (Durlak, et al., 2011).

**Summary of Previous Syntheses**

Mounting evidence suggests mental health prevention programming may be an effective means of promoting school performance among children and adolescents. However, few studies beyond school-based SEL programs have capitalized on this approach to promoting both youth well-being and educational outcomes (Becker & Luthar, 2002). Previous reviews and meta-analyses have provided evidence to suggest that mental health prevention efforts may be efficacious or effective in promoting positive school outcomes among children and adolescents. However, none of these studies have fully filled this gap in the literature because they have been limited in important ways. First, all universal and selected PMH focused studies have been exclusive to examining whether mental health programming is efficacious or effective in school-based settings only. There is some evidence to suggest that indicated PMH, in the form of SEL,
may be effective outside of school. However, some SEL programming differs conceptually from general PMH. Thus, it is yet to be known whether universal and indicated PMH programming may affect school outcomes when implemented outside of the school setting. Second, these previous studies were very limited in the types of school outcomes examined. Performance-related outcomes were the only school outcomes analyzed. Authors have suggested that examining only performance outcomes in the educational realm may present a limited picture of the full impact of mental health programs on school outcomes (Durlak, et al, 2011), such as academic self-efficacy or sense of school belonging. Other limitations specific to each study include questionable methodology (e.g., Hoagwood, et al., 2007), excluding prevention studies (e.g., Baskin et al., 2010), and focusing solely on SEL prevention (e.g., Durlak et al., 2011). The current meta-analysis will address each of these limitations.

The exclusion of examining school outcomes in mental health prevention intervention is problematic. One way to illustrate this problem is to observe the need versus availability of school-based mental health services in the U.S. School-based mental health is essentially the provision of counseling or psychological services within the school setting by trained professionals (Foster, Rollefson, Doksum, Noonan, & Robinson, 2005). School-based mental health provides an accessible means of seeking services. However, the proportion of students in need of services continues to outnumber available resources (Hoagwood et al., 2007). A primary explanation for the lack of resources allocated to mental health services in schools is that the benefits of mental health
services are often divorced from school benefits (Hoagwood et al., 2007). The majority of studies of mental health interventions fail to include basic measures of school-related outcomes (Hoagwood & Johnson, 2003). As a consequence, the impact of mental health interventions on both mental health and educationally relevant behaviors is poorly understood (Hoagwood et al., 2007). For example, Weisz et al. (1995) stated that he and his research team chose to exclude school outcomes from a meta-analysis examining the effectiveness of school-based mental health interventions because “so many factors (e.g., intelligence) other than psychopathology could be responsible for poor academic performance that it seemed inappropriate to base tests of psychotherapy efficacy on such outcomes” (p. 455). Although Weisz and his colleagues are correct in their conclusion that school outcomes are likely influenced by a constellation of variables that could overshadow the unique contribution of mental health functioning, we cannot ignore the potential contribution of mental health on school functioning.

**Moderators of Program Effectiveness**

Although the primary focus of the current study was to determine whether mental health prevention targeting youth positively affects school outcomes, this study also examined moderators of program outcomes to allow for between group differences. Thus, questions regarding the circumstances in which these programs are most beneficial were answered. Previous research has suggested that the following factors are important to consider (Becker & Luthar, 2005; Durlak & Wells, 1997, 1998; Durlak et al., 2011, Ellis, 1998; Greenberg, Domitrovich, and Bumbarger, 2000): a) multi-systemic prevention programming, b) youths’ age, c)
youths’ socio-economic status (SES), d) program length, e) prevention level, and f) program setting.

**Multi-Systemic Prevention Programming**

As mentioned above, Ellis (1998) contended that for prevention efforts to maximize their benefit, they must be multi-systemic. From an ecological perspective (Bronfenbrenner, 1994), youth are embedded within a set of systems and settings that differ in size and proximity of influence. Multi-systemic intervention programming may manifest in a number of different forms. Utilizing multiple settings or domains in which the child interacts or focusing intervention efforts on both the child and one or more person(s) within their microsystem all constitute multi-systemic intervention. Often these programs are referred to as multi-systemic, multi-component, or multi-domain interventions.

Several studies have provided evidence that multi-systemic mental health–related interventions are more effective approaches to prevention than single-component interventions. For example, Catalano and colleagues (2005) examined 25 programs incorporating positive youth development constructs into universal or indicative approaches with youth between the ages of 6 and 20 years; the most effective programs involved multiple domains of the child’s microsystem, such as involving individuals in schools, churches, or community agencies. Two single social domain community interventions and six school interventions were identified as effective. Seven school-family domain interventions and one community-school intervention were identified as effective. Among interventions involving three domains, seven family, school, community interventions, one
family, church, and community intervention, and one community, school, work-setting intervention were identified as effective.

Hoagwood et al. (2007) and Durlak et al. (2011) noted that many school-based studies included in their reports were multi-systemic, in that they involved not only the child, but for instance, also intervened with the parents or teachers. Hoagwood et al. (2007) noted that the studies that yielded more significant results were interventions that were multi-systemic, involving multiple components that included other agents of the child’s microsystem. Again, Hoagwood and colleagues’s (2007) results should be interpreted with caution, given that the review did not empirically test the overall efficacy of these studies. Providing further evidence, Durlak et al. (2011) noted in their meta-analysis that the most efficacious school-based SEL interventions were multi-systemic.

Researchers (i.e., Hoagwood et al., 2007; Durlak et al., 2011) suggested that future studies should examine in further detail the characteristics of these PMH programs that are multi-systemic. No meta-analytic studies, to date, have examined whether multi-systemic mental health prevention programs are more effective at increasing school outcomes among youth, than single-component programs.

Age

Mental health prevention programs for youth spans from childhood to late adolescence. The current study will limit the inclusion of studies to youth of school-age (kindergarten to 12th grade), approximately age 4 to 18 years of age. This age range remains broad and spans across distinctively different
developmental levels. Previous literature has noted differences in effectiveness of mental health interventions across age groups (Greenberg, Domitrovich, & Bumbarger, 2000). Through their thorough review of the literature, Greenberg et al. (2000) concluded that PMH efforts may be most useful when implemented early. Authors cited the resistance to treatment that serious conduct disorders may have when not addressed in childhood (Greenberg et al., 2000). However, evidence does suggest that middle school adolescents may yield even greater benefits from prevention efforts because it is within these years, adolescents’ teacher, classroom, and school experiences have long-lasting effects on future educational and life opportunities (Eccles & Wigfield, 1997).

**Youth Socio-Economic Background**

Social class and ethnicity have often been confounded in the intervention literature. Often “disadvantaged” or “at-risk” are used as a term that refers to non-White individuals (e.g., Becker & Luthar, 2002). An explanation for this phenomenon is that there is a disproportionate number of ethnic minorities in the United States that are economically disadvantaged. However, it is important to note that disadvantaged or at-risk do not always equate to ethnic minority. This becomes of particular importance when attempting to analyze background characteristics of youth within studies. For the purposes of the current study, youth demographics will be examined as a potential moderator. These demographic characteristics refer to race/ethnicity and socio-economic status (SES). As previously mentioned, mental health may be especially important for
the academic success of low-income, ethnic minority youth (Becker & Luthar, 2002).

**Program Dosage**

Dosage represents the amount of prevention sessions administered to participant(s). For multi-systemic interventions, dosage extends to the intervention provided to other agents involved in the intervention. For instance, in an intervention that has children who participate in 15 sessions of skills training and parents who participate in parent training for five sessions would have a dosage of 20. A meta-analysis of school-based SEL noted that the most efficacious universal programs had greater dosages than those that were less efficacious (Durlak et al., 2011). Evidence suggests that short-term PMH programs produce only time-limited benefits, if at all, particularly with at-risk groups, whereas multi-year programs are more likely to produce enduring benefits (Greenberg, Domitrovich, & Bumbarger, 2000).

Even though dosage may be a particularly important moderator to consider, it may confound intervention fidelity. Programs that have a greater dosage are often times more difficult for participants to adhere to, given the greater time duration (Durlak & DuPre, 2008). Dosages that equate to longer lengths of time are a common reason why many interventions that utilize parents often report less than effective outcomes (Greenberg, Domitrovich, & Bumbarger, 2000). Parents, especially in lower SES communities, are often more difficult to engage in the full dosage that the intervention requires to reach its full potential of effectiveness (Durlak & Wells, 1997). Analyzing the effects of dosage on school
outcomes may help reveal whether poor effect sizes may be a result of poor program fidelity related to dosage, rather than general program ineffectiveness.

**Prevention Level**

All studies to be included in the current investigation will be preventive in nature. That is, they will include youth who have yet to be identified as having a psychological disorder, nor will they have been classified as failing in school. However, as mentioned previously, three prevention levels exist that will be included in this investigation – universal, selected and indicative. Some meta-analyses of youth mental health interventions have noted that indicative prevention may yield stronger effects than universal prevention (Greenberg et al., 2000), while others have evidenced that universal and more selected approaches are similarly effective at reducing mental health problems among youth (Durlak & Wells, 1997, 1998). However, it is unclear whether these findings hold for school outcomes. Perhaps higher functioning youth would benefit more than an indicative sample because they will not have the same barrier of overcoming current struggles in order to reap the secondary benefits of mental health programming.

**Intervention Setting**

Finally, intervention setting will be included in analysis. This moderator is of utmost importance because most studies that have examined the role of mental health prevention on school outcomes have been exclusively school-based (e.g., Hoagwood et al., 2007; Durlak et al., 2011). Thus, it is unclear whether mental health prevention may affect school outcomes if implemented in, for
instance, a clinic-based setting. Theoretically, the same skills will be taught in a clinic-based program as a school-based program. Baskin et al. (2011) included psychotherapy treatment programs that spanned across various settings. However, authors did not analyze whether setting affected school outcomes. Therefore, the question remains whether setting matters. Based on the importance of involving multiple systems in prevention efforts, Ellis (1998) suggested that intervention is most effective when implemented within the settings that the child interacts with the most. In fact, Greenberg and colleagues (2000) explicitly stated that for children and adolescents, the school context should be a “central focus of intervention (p. 38).”

Rationale for the Current Investigation

The U.S. ranks in the bottom third of all industrialized countries in math and science (Hanushek et al., 2010). Educational disparities are greater and even more widespread among low-income, ethnic minority students (Becker & Luthar, 2005). Recently, policy-makers and researchers have begun to focus on the role of the overall mental health of the child in school outcomes. Indeed, numerous studies have documented the association between mental health and educational outcomes, noting that the two are interrelated (e.g., Roeser et al., 1999; Suldo & Shaffer, 2008). However, the link between mental health and school outcomes continues to be understudied, particularly in the intervention arena.

Universal, selected, and indicated PMH programs are well-established means of reducing the incidence of mental illness (Weisz, Sandler, Durlak, & Anton, 2005). Meta-analytic studies have confirmed the effectiveness of these
prevention approaches to promote mental health through reducing problem behaviors and promoting key competencies (Durlak & Wells, 1997, 1998). Due to the often multi-systemic and multi-factor (e.g., risk and protective factors) approach of PMH, researchers have suggested that PMH efforts may yield benefits that extend to other domains of youth functioning (Greenberg, Domitrovich, & Bumbarger, 2000). However, no meta-analysis to date has examined whether PMH is an intervention approach to promote positive school outcomes among youth. Past meta-analyses and reviews that have attempted to examine this relationship have had limitations in scope or in methodology. The current investigation sought to bridge the gap in knowledge about the effects of PMH on a range of school outcomes by using meta-analytic strategies to pool all studies within the selected eligibility criteria in order to synthesize quantitative study outcomes.

Finally, because it has been well-documented that universal and indicated PMH programs are overall effective at promoting positive mental health outcomes, the current meta-analysis did not seek to duplicate existing literature by reporting mental health outcomes. The primary focus of this study was to report academic achievement (e.g., grades, standardized test scores, attainment) and educationally relevant outcomes (e.g., academic self-efficacy, attendance, school engagement) associated with PMH interventions. This goal is particularly important because few previous syntheses examining PMH programs’ impact on school outcomes have limited their focus to academic performance indicators only.
Research Questions and Hypotheses

The overarching goal of the current investigation was to meta-analyze studies of PMH programs in order to evaluate their ability to produce educational benefits for youth. As a secondary goal, but equally important goal, this meta-analysis examined factors, or moderators, that influenced program effectiveness.

Meta-analysis was used to test the following hypotheses. Specifically, tests of pre-post effect sizes were evaluated through meta-analysis to address the overarching question of program benefits. Analyses then examined whether effect size variation was associated with differences in specific moderators (Cooper, 2010).

**Hypothesis I.** As suggested by previous meta-analyses and research reviews (Durlak & Wells, 1997,1998; Durlak et al., 2011; Hoagwood et al., 2007) it was predicted that youth PMH programs, overall, would increase school outcomes among youth. This effect size was expected to be in the small range.

**Hypothesis Ib.** It was predicted that school outcomes that represent more proximal indicators of the change process (e.g., attitudes) would yield a greater effect size than those that represent more distal indicators of change (e.g., behavioral, performance).

**Hypothesis IIa.** Programs that are multi-systemic, that is, those that involve intervention at other levels of the child’s micro- or meso-system (e.g., parents, teachers, school-wide reform), would be more effective than programs that are single component.
**Hypothesis IIb.** Educational benefits of mental health programming would decrease as age increases, peaking in pre-adolescent years.

**Hypothesis IIc.** There would be no difference in educational outcomes based on youths’ ethnic/racial background. However, it was expected that youth from more socio-economically disadvantaged backgrounds would exhibit greater educational gains than those from more advantaged backgrounds.

**Hypothesis IIc.** There would be no difference in educational outcomes based on youths’ ethnic/racial background. However, it was expected that youth from more socio-economically disadvantaged backgrounds would exhibit greater educational gains than those from more advantaged backgrounds.

**Hypothesis IIId.** Programs with a larger dosage would be positively associated with greater educational benefits. However, it was predicted that the educational benefits associated with dosage would reach a ceiling, and benefits would be unaffected by dosage after a certain amount.

**Hypothesis IIe.** It was predicted that indicated interventions would be more effective at producing positive educational outcomes for youth, compared to universal prevention interventions.

**Hypothesis IIf.** It was expected that program setting would moderate program effectiveness. In general, it was expected that programs that were implemented in settings that the individual child interacts with the most would be more effective than environments that the child has comparatively less contact. Given the contextual relevance of the school setting to school outcomes and the salience of this domain for most youth, programs incorporated in the school setting would be most effective.
CHAPTER II

METHOD

In order to assess the effectiveness of PMH interventions on school outcomes among children and adolescents, a meta-analysis of quantitative evaluations was conducted to answer the research questions and test study hypotheses. Research Hypotheses 1 and 2 were tested by meta-analytic strategies. Meta-analysis includes these steps: (a) defining study eligibility criteria, (b) searching and locating eligible studies, (c) coding study characteristics and using available statistical information to compute effect sizes, (d) calculating an overall/average effect size comprised of findings from all studies as well as an estimate of the degree to which effect size varies across studies, and (e) assuming there is significant variation in effect sizes, conducting moderator analyses to examine study characteristics that may be associated with and thus account for this variation (Cooper, 2010; Lipsey & Wilson, 2001).

This chapter discusses the first two steps. Study eligibility was guided by Lipsey and Wilson’s (2001) suggestions to aid the meta-analyst in organizing and defining study criteria. Chapter III addresses the remaining three steps of meta-analysis.

Inclusion/Exclusion Criteria

First, to be included in this meta-analysis, studies had to meet the requirements pertaining to distinguishing features. The primary distinguishing features for inclusion in this analysis was that the study must involve the evaluation of a preventive mental health intervention for youth. Generally, the
distinguishing features of PMH were guided by Durlak & Wells’ seminal articles (1997; 1998). PMH programs were defined as, “an intervention intentionally designed to reduce the future incidence of adjustment problems in currently normal populations as well as efforts directed at the promotion of mental health functioning” (Durlak & Wells, 1997, p. 117). Prevention included either universal, selected, or indicated levels. Universal prevention approaches were identified when “all members in an available population receive[d] the intervention” (Durlak & Wells, 1997, p. 118). Selected prevention programs were those that selected “groups considered at risk for eventual problems, but who are not yet dysfunctional” (Durlak & Wells, 1997, p. 118). Indicated PMH programs were defined as, “interventions that seek to identify early signs of maladjustment and to intervene before full-blown disorders develop” (Durlak & Wells, 1998; p. 776).

Second, in terms of research respondents, the study had to include a sample of school-age youth, operationalized as individuals between the ages of 4 and 18 years of age or in kindergarten through 12th grade. Youth too young or old to be in school were excluded because the target of this meta-analysis was on school outcomes associated with academic functioning of youth through high school.

Next, studies had to meet requirements related to key variables. First, included studies reported at least one quantitative school outcome. School outcomes were defined as outcomes directly relevant to youth’s current or future school functioning (Baskin et al., 2010; Becker et al., 2014). These school
outcome variables included traditional achievement (e.g., grades, test scores, grade promotion) and behavioral (e.g., misconduct, attendance) indicators. The aforementioned outcomes are typically considered more distal change outcomes (Durlak et al., 2011, Fredrickson et al., 2006). As suggested in previous studies (e.g., Durlak et al., 2011), the current meta-analysis also examined more proximal change indicators relevant to school functioning (e.g., school engagement, academic motivation, academic self-efficacy, attitudes about school). Second, included studies provided sufficient data to calculate the effect size. Various strategies were used to calculate effect sizes from reported data, these strategies are discussed further in the Data Analysis section. When information that is required to compute an effect size was missing from an article, attempts were made to obtain such data from the study authors.

Fourth, studies in this meta-analysis had to meet the following research methods requirements: (a) the study must include a control or comparison group. The control conditions included were “treatment as usual,” placebo, wait-list, or no treatment. The key was that the control condition represented youth who did not receive PMH services. Nonequivalent comparison designs in which groups were not randomly assigned to conditions were eligible only if there were pre-test measures on mental health and school outcomes. Pre-test measures help to identify threats to internal validity associated with non-equivalent comparison designs. One group pre- post-test designs were not eligible; (b) included studies were from independent samples. Specifically, studies that used data from the same sample were included to the extent that they differ in outcomes and/or
moderators analyzed. Multiple studies that reported data from the same sample were not included more than once in the analysis of an overall effect size.

In terms of language, to be included, all studies had to be reported in English. However, studies conducted in countries outside of the U.S. were reviewed.

Publication types included peer-reviewed published studies and unpublished studies in the form of dissertations, theses, and other manuscripts that were not published due to null findings. Meta-analyses that exclude unpublished works present a methodological flaw that may result in an upward bias in the size of effects that are found (Cooper, 2010; Lipsey & Wilson, 2001). Methods for testing publication bias are discussed in the data analytic section below. Technical reports and conference presentations were excluded.

Finally, the time frame for study inclusion was from 1980 to 2012. This time frame limitation is important because factors associated with less recent times may present significant contextual and historical confounds (Lipsey & Wilson, 2001).

**Literature Search Procedures**

A research team consisting of several undergraduate-level research assistants and the first author searched the literature for eligible studies from January to May 2013. Guided by the above inclusion/exclusion criteria, searching for eligible studies began with a review of the following major online database search engines: PsychInfo, Academic Search Premier, ERIC, PubMed, and Social Science Citation Index, Google Scholar, Proquest Dissertations and Theses
Database. Keywords, in part informed by previous studies (Baskin et al., 2011; Hoagwood et al., 2007), were used to search for relevant studies. Keywords included combinations of four domains. First, keywords related to the mental health domain included, “mental health,” “emotional/behavioral problems,” “psychiatric,” “therapy,” “well-being,” and “positive youth development.” Second, school-focused terms included, “academics,” “school/educational/academic outcomes,” and “achievement.” Third, keywords that captured the population were, “children,” “adolescents,” and “youth.” Finally, keywords pertinent to the intervention component included, “prevention,” “intervention,” “evaluation,” “outcome,” and “study.”

Next, eligible studies were sought by examining the reference sections of published reviews and meta-analyses that included studies with PMH samples of youth, while including school outcomes (Baskin et al., 2011; Durlak et al., 2011; Hoagwood et al., 2007). Other journal articles and scholarly handbooks pertaining to mental health and school outcomes (e.g., Psychological Counseling Research Focus), were searched to determine whether there were other reviews or meta-analyses that may have cited target studies.

As a final step after identifying as many eligible studies as possible through the traditional channels, efforts were made to contact researchers to obtain unpublished studies. Attempts were made to contact authors who have published two or more articles related to the current field of study in order to ask for unpublished studies (e.g., Drs. Joseph Durlak, John Weisz, Howard Adelman, Irwin Sandler, Gerald August, etc.).
Throughout the search process, the research team met weekly to discuss ambiguous eligibility concerns. Most commonly, the question emerged whether a program was truly a PMH program or some other general positive youth development/health/education program. For example, studies that sought to promote parent engagement in the school setting were excluded, as well as general extracurricular activity youth development programs (e.g., an after-school program promoting participation in a team sport). In cases in which the research team was unsure of study inclusion after discussing together, the first author sought consultation with another PhD level clinical-community psychologist.

**Coding Procedure**

Coding is a process that involves “interviewing” (p. 73) eligible studies in order to answer specific questions of interest to the meta-analyst (Lipsey & Wilson, 2001). As described by Lipsey and Wilson (2001), coding encompasses two major categories of information: details about study characteristics, or study descriptors, and details about empirical findings of the study, or effect sizes. These two categories generally can be thought of as encoding information relevant to independent (study descriptors) and dependent (effect sizes) variables. Study descriptors that were coded for in each study included source information (e.g., publication form and year), methods (e.g., sample descriptors, methodological design), program characteristics (dosage, mode of delivery, programmatic components), other independent variables relevant for moderation analyses (e.g., program setting, multi-systemic or individual-focused programs), statistical data (e.g., appropriate means, standard deviations), and related outcome
variable information. Appendix A includes the coding manual. Given the
iterative nature of the coding process, the coding manual was revised as studies
were reviewed and additional variables of interest emerged (Lipsey & Wilson,
2001). For instance, program length was added into the coding manual because it
became clear early in the coding process that indicators of program dosage were
not consistently reported.

Study coding was conducted with the aid of two teams of undergraduate
assistants. The undergraduate assistants were oriented to and trained on the coding
manual. The first author trained one team of undergraduates to code study
characteristics between January to June 2013, while a PhD level researcher who
had conducted past meta-analyses trained the second team of research assistants
on coding effect size data between September to December 2013. Once the coders
were familiar with the coding protocol, several studies were selected for each of
the coders to practice coding. Coders compared their resulting coded studies with
one another and with the author. Inter-rater reliability was assessed and each
coder was able to code independently after kappa of 80% was reached for at least
two study codings. Kappa represents the number of agreements in coding divided
by the number of observations and then multiplied by 100 to represent the
percentage of inter-rater agreement (Cohen, 1968). As an additional method of
ensuring reliable coding, all studies were coded in pairs in which one other
individual coded the study for coding replication. This approach allowed for an
on-going check of the quality of coding reliability and helped to identify any areas
in which coders may have had difficulty (Lipsey & Wilson, 2001). Finally,
throughout the coding process, the undergraduate assistants met weekly with the author and the PhD-level co-investigator to address problems or areas of confusion and clarify coding procedures. When cases presented with insufficient study information necessary to compute an effect size or other essential information was missing, the corresponding study authors were contacted to obtain this information.
Computing Effect Size

Research findings were reported in a variety of formats. The most common was the pre-post group contrasts, in which two or more groups of respondents are measured on a variable and then those responses are compared (Lipsey & Wilson, 2001). Additionally, given that school outcomes were measured a number of different ways (e.g., GPA, engagement, academic self-efficacy) across numerous studies, pre-post group contrast means were standardized (Baskin et al., 2011). Thus, the standardized mean difference was commonly used to compute the effect sizes, or Cohen’s $d$, in the current investigation. The formula is:

$$d = \frac{M_1^{Adjusted} - M_2^{Adjusted}}{s_{Pooled}},$$

where $M_1^{Adjusted}$ and $M_2^{Adjusted}$ are the sample means of the two independent groups accounting for the correlation between pre- and post-test. $s_{Pooled}$ is the within-groups standard deviation, pooled across groups (Cooper et al., 2009). Various estimation procedures were employed depending on whether all necessary data (e.g., means, standard deviations, test statistics) were available to accurately compute effect sizes (Lipsey & Wilson, 2001). For instance, if pre-
and post-test data were not provided but F-tests or T-tests were reported, the appropriate effect size was estimated from these values (Lipsey & Wilson, 2001).

Hedges $g$ was calculated using the statistical software Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2007) in order to correct for biases associated with sample size. Because effect size formulas compute the magnitude of an effect, independent of sample size, studies with smaller samples sizes were computed as being equal in their overall contribution as a study with a very large sample size. This is problematic because smaller sample sizes are inherently less precise than larger sample sizes and should not be considered as being equally reliable. Hedges $g$ (Hedges & Olkin, 1985) provides an approach to weight each effect size value to represent its relative precision. In this formula Cohen’s $d$ is converted to Hedges $g$ by weighting each effect size value by its sample size, called the inverse variance weight. To explore the potential effects of publication bias (Hunter & Schmidt, 2004), I used a number of methods using CMA (Borenstein, Hedges, Higgins, & Rothstein, 2007). Each of these tests of publication bias are described in the Results Section.

**Analysis of Overall Program Effectiveness**

The next step in analysis involved determining the unit of analysis and the statistical model (Cooper et al., 2009). This meta-analysis used the independent sample as the primary unit of analysis. When effect size data were reported for the overall study sample then each individual study contributed one sample to the analysis. In cases in which effect size data were reported for subgroups within the overall study’s sample (e.g., subgroups based on race/ethnicity) each subgroup
sample acted as an independent sample (Cooper et al., 2009). Each independent sample contributed one effect size to the calculation of a mean/overall effect size measure across outcomes. Numerous past meta-analyses of mental health interventions have reported on effect sizes of psychological variables and found significant positive effect sizes (e.g., Durlak & Wells, 1997; 1998; Farahmand et al., 2011; Farahmand et al., 2012; Prout & Prout, 1998). Because the current investigation is on determining whether mental health programs affect school outcomes, only outcome variables related to school functioning were considered for effect size calculation.

A random effects model was assumed for the current investigation. A random effects model, rather than a fixed effects model, accounts for both random sampling error and the variance associated with random study-level difference (Lipsey & Wilson, 2001). Thus, random effects models are considered to be more conservative than fixed effects models (Cooper et al., 2009). Several authors suggest assuming a random effects model for analysis because of the assumed variance one should expect in numerous studies’ procedures, sample, settings, and other study-specific characteristics (Hedges & Vevea, 1998; Lipsey & Wilson, 2001). Studies included in this meta-analysis were highly heterogeneous. Studies varied in a number of ways, such as in the characteristics of the youth involved in the mental health program, program design, and specific outcomes measured. These sources of variability warranted the use of a random effects model for analysis.
Using a random effects model, study-level variance of mean effect sizes was computed and tested for significant differences between studies. The null hypothesis in this analysis assumes no significant differences exist between studies. The test statistic in this analysis is called $Q$, and is often called a test of homogeneity (Lipsey & Wilson). If $Q$ is significant (rejecting the null hypothesis) heterogeneity across studies is assumed further analyses (i.e., moderator) will be conducted in an effort to explain the sources of variance across study mean effect sizes. Informed by Baskin and colleagues’ (2011) and Becker (2014), it was predicted that analyses would yield significant heterogeneity based on the expected high variation in types of school outcome measures.

Finally, an overall weighted standardized mean effect size ($g$) across all studies and its 95% confidence band were computed. Additionally, $gs$ and 95% confidence intervals were computed for each outcome category. For the overall analysis, this investigation aggregated outcome scores so that each study contributed only one effect size. However, in order to address Hypothesis 1b, effect sizes for outcome categories were computed separately. That is, separate overall outcome analyses were run for performance, discipline, classroom behavior, and academic attitude outcome variables.

**Moderator Analyses**

Finally, after overall effect sizes were calculated to determine the effect of PMH programs on school outcomes, moderation analyses were conducted to determine how these specific factors influenced effect sizes. Moderators are listed in Hypothesis II.
Lipsey and Wilson (2001) cautioned that moderators should only be analyzed if they are representative of a large enough amount of studies and if there is significant unexplained variance in effect sizes. Moderator variables were analyzed if they were characteristic of a large enough number of studies and if there was significant heterogeneity yielded from meta-analytic data analysis. Continuous and categorical moderators were analyzed separately. Categorical moderator variables were dummy-coded tested for significant differences between groups using the Q statistic. Continuous moderator variables were included as predictors in meta-regression models. All moderator analyses were conducted using CMA (Borenstein, Hedges, Higgins, & Rothstein, 2007).
CHAPTER IV

RESULTS

Search Outcome

Using the abovementioned search methods, over 1,400 articles were yielded. An examination of titles and abstracts limited the search results to 142 studies. Approximately two-thirds of studies were yielded from reference checks of previous syntheses reporting findings related to school outcomes (e.g., Baskin et al., 2010; Durlak et al., 2011; Hoagwood et al., 2007; Payton et al., 2009; Sklad et al., 2010). The first author examined each of the 142 studies in detail to determine whether the study fit all eligibility criteria. Of the 142 studies, 34 studies with 46 independent program and control/comparison samples (e.g., low- and high-risk subgroups, males and females) were included in this meta-analysis. Studies were excluded mainly due to the following: no school outcome data reported, more than 50% of youth participants met criteria for a DSM-IV mental disorder, the prevention program was not primarily focused on mental health promotion or mental illness reduction (e.g., parental school involvement intervention, childhood literacy promotion), study design did not meet criteria (e.g., no control group). Table 1 presents key characteristics of included studies.
Table 1

*Descriptive Characteristics of 34 PMH Interventions*

<table>
<thead>
<tr>
<th>Participant Characteristics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Child age in years</td>
<td>$M = 11.25^*$</td>
</tr>
<tr>
<td></td>
<td>(SD = 3.41)</td>
</tr>
<tr>
<td>Treatment samples size</td>
<td>$M = 351.46$</td>
</tr>
<tr>
<td></td>
<td>(SD = 781.39)</td>
</tr>
<tr>
<td>Comparison samples size</td>
<td>$M = 241.06$</td>
</tr>
<tr>
<td></td>
<td>(SD = 436.07)</td>
</tr>
<tr>
<td>Socio-economic status</td>
<td>&gt; 60% economically disadvantaged</td>
</tr>
<tr>
<td></td>
<td>60%</td>
</tr>
<tr>
<td>Ethnic Diversity</td>
<td>&gt; 60% White</td>
</tr>
<tr>
<td></td>
<td>39.40%</td>
</tr>
<tr>
<td></td>
<td>&gt; 60% Minority</td>
</tr>
<tr>
<td></td>
<td>36.40%</td>
</tr>
<tr>
<td></td>
<td>&gt; 60% Neither White nor Minority</td>
</tr>
<tr>
<td></td>
<td>24.20%</td>
</tr>
<tr>
<td>Educational Level</td>
<td>Elementary (K-6)</td>
</tr>
<tr>
<td></td>
<td>51.50%</td>
</tr>
<tr>
<td></td>
<td>Middle School (7-8)</td>
</tr>
<tr>
<td></td>
<td>18.20%</td>
</tr>
<tr>
<td></td>
<td>High School (9-12)</td>
</tr>
<tr>
<td></td>
<td>18.20%</td>
</tr>
<tr>
<td></td>
<td>Mixed (4-9)</td>
</tr>
<tr>
<td></td>
<td>12.10%</td>
</tr>
<tr>
<td>Manuscript Year</td>
<td>1980 - 1989</td>
</tr>
<tr>
<td></td>
<td>8.60%</td>
</tr>
<tr>
<td></td>
<td>1990 - 1999</td>
</tr>
<tr>
<td></td>
<td>34.30%</td>
</tr>
<tr>
<td></td>
<td>2000 - 2012</td>
</tr>
<tr>
<td></td>
<td>57.10%</td>
</tr>
<tr>
<td>Published studies</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>85.70%</td>
</tr>
<tr>
<td>Study design</td>
<td>Included some randomization</td>
</tr>
<tr>
<td></td>
<td>71.40%</td>
</tr>
<tr>
<td>Implementation</td>
<td>Research Team</td>
</tr>
<tr>
<td></td>
<td>32.40%</td>
</tr>
<tr>
<td></td>
<td>School/Community Personnel</td>
</tr>
<tr>
<td></td>
<td>61.80%</td>
</tr>
<tr>
<td></td>
<td>Research Team and Community Personnel</td>
</tr>
<tr>
<td></td>
<td>5.90%</td>
</tr>
<tr>
<td>Multi-systemic</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>54.34%</td>
</tr>
<tr>
<td>Setting</td>
<td>School</td>
</tr>
<tr>
<td></td>
<td>85.70%</td>
</tr>
<tr>
<td></td>
<td>Broader Community</td>
</tr>
<tr>
<td></td>
<td>8.57%</td>
</tr>
<tr>
<td>Clinic</td>
<td>5.70%</td>
</tr>
</tbody>
</table>

*More than 50% of studies did not report age of child participants*
PMH Effects on School Outcomes

Data were assessed for 34 studies across 46 independent samples. Using a random effects model, the overall effect size (reported in Hedge’s g) calculated across coded school outcomes, yielded an effect size of .201, with a 95% confidence interval (CI) of .121 to .282. To clarify, all individual effect sizes were coded as positive when the outcome favored the intervention group. Further, only post-test outcomes were assessed for all analyses because over 70% of studies did not report outcomes at follow-up. Borenstein and colleagues (2009) recommend that outcomes should only be meta-analyzed when a sufficient proportion of the data are reported to represent the effect. The effect size of .201 was significantly different from zero (p < .001). Thus, the null hypothesis was rejected suggesting that PMH programs have a small, positive effect on school outcomes for youth who receive PMH programming compared to youth who do not receive PMH programming. A table summarizing each study’s characteristics and school outcomes measured is presented in Table 2.
Table 2

*Study Characteristics*

<table>
<thead>
<tr>
<th>First Author &amp; Year</th>
<th>Target Problem</th>
<th>Study Design</th>
<th>Child Sample Size</th>
<th>Delivery</th>
<th>Intervention</th>
<th>School Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbuthnot (1986)</td>
<td>Delinquency</td>
<td>Youth matched and randomly assignment</td>
<td>24 program, 24 comparison</td>
<td>Research team</td>
<td>Indicated, school-based, small group moral reasoning development</td>
<td>GPA, Absenteeism, Tardiness</td>
</tr>
<tr>
<td>August (2001)</td>
<td>Aggressive behaviors</td>
<td>Random assignment of 10 program and 10 control schools</td>
<td>124 program, 121 comparison</td>
<td>Trained community staff members.</td>
<td>Indicated, community-based summer school, ongoing teacher consultation and student mentoring, bi-weekly parent training, parent support</td>
<td>Academic competence, discipline</td>
</tr>
<tr>
<td>Brackett (2010)</td>
<td>General socio-emotional development</td>
<td>Quasi-experimental, randomization no matching</td>
<td>155 program, 118 comparison</td>
<td>Teachers</td>
<td>Universal, classroom-based social-emotional learning curriculum</td>
<td>Teacher-rated class behavior, math and English grades</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>Method</td>
<td>Group 1</td>
<td>Group 2</td>
<td>Outcomes</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Burdsal (1980)</td>
<td>Delinquency</td>
<td>Quasi-experimental</td>
<td>46 program, 21 comparison</td>
<td>Trained community staff members</td>
<td>Indicated, community-based day camp for youth, family therapy, and school consultation</td>
<td></td>
</tr>
<tr>
<td>Cappella (2012)</td>
<td>Antisocial behaviors</td>
<td>Teachers randomly assigned within schools</td>
<td>74 program, 85 comparison</td>
<td>Mental health consultants &quot;indigenous&quot; to the community</td>
<td>Universal, school-based teacher training and coaching in classroom management</td>
<td></td>
</tr>
<tr>
<td>Castro-Villareal (2009)</td>
<td>Substance use</td>
<td>Quasi-experimental pre- and post test. RY (n = 70), comparison (n = 30)</td>
<td>70 program, 30 comparison</td>
<td>Teachers</td>
<td>Indicated, classroom-based course in developmental competencies</td>
<td></td>
</tr>
<tr>
<td>Catalano (2003)</td>
<td>Antisocial behaviors and promotion of pro-social behaviors</td>
<td>Schools randomly assigned to treatment or TAU</td>
<td>497 program, 441 comparison</td>
<td>School-home coordinators and staff development coordinator</td>
<td>Universal, teacher workshops for classroom management, instructional strategies. Parent training workshops and student summer camp</td>
<td>Attendance, class behavior, objective discipline, Academic self-concept, Attendance Grades, Commitment to school, academic performance,</td>
</tr>
<tr>
<td>Study</td>
<td>Outcome</td>
<td>Design</td>
<td>Description</td>
<td>Intervention</td>
<td>Control</td>
<td>Outcomes</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td>-------------</td>
<td>--------------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Coleman (2000)</td>
<td>Externalizing behaviors</td>
<td>Randomized controlled trial</td>
<td>7 school matched and randomly assigned to CCS training or comparison condition.</td>
<td>Research team</td>
<td>Universal, school-based classroom management intervention taught to teachers</td>
<td>Suspensions</td>
</tr>
<tr>
<td>Eggert (1993)</td>
<td>Substance use and school drop-out</td>
<td>Quasi-experimental design</td>
<td>Four schools in one district, at-risk youth randomly assigned to condition.</td>
<td>Teachers</td>
<td>Selected, school-based &quot;Personal Growth&quot; course involving group support and life-skills training.</td>
<td>Perceived school bonding, GPA, attendance</td>
</tr>
<tr>
<td>Flay (2001)</td>
<td>Character development and problem behavior prevention</td>
<td>Schools in Nevada and Hawaii, matched by demographics</td>
<td>20 program school, 40 comparison schools</td>
<td>School personnel</td>
<td>Universal, school-climate change, teacher training, family/parent literacy and skills, child curriculum focused on pro-social/character development</td>
<td>Achievement, suspension, attendance</td>
</tr>
<tr>
<td><strong>Gottfredson (1993)</strong></td>
<td><strong>Disruptive behavior</strong></td>
<td><strong>Non-equivalent control group design</strong></td>
<td><strong>4064 program, 1214 comparison</strong></td>
<td><strong>School personnel</strong></td>
<td><strong>Universal, school-wide discipline policy change, behavior tracking of students, classroom management, positive reinforcement system</strong></td>
<td><strong>Disruptive and on-task behavior in class</strong></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------</td>
<td>----------------------------------------</td>
<td>--------------------------------</td>
<td>---------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td><strong>Gottfredson (2002)</strong></td>
<td><strong>Externalizing behaviors</strong></td>
<td><strong>Quasi-experimental - No randomization to condition</strong></td>
<td><strong>97 program, 74 comparison</strong></td>
<td><strong>Research team</strong></td>
<td><strong>Universal, school-based group-based CBT social skills and problem-solving</strong></td>
<td><strong>GPA, Absent, Tardy, Suspensions</strong></td>
</tr>
<tr>
<td><strong>Hains (1994)</strong></td>
<td><strong>Stress inoculation</strong></td>
<td><strong>Randomly assigned to program group and waitlist control group, not matched for equivalence</strong></td>
<td><strong>10 program, 11 comparison</strong></td>
<td><strong>Research team</strong></td>
<td><strong>Universal, school-based group and individual sessions - cognitive restructuring, problem-solving, anxiety management training</strong></td>
<td><strong>Absenteism, GPA</strong></td>
</tr>
<tr>
<td>Study</td>
<td>Outcome</td>
<td>Random Assignment</td>
<td>Sample Size</td>
<td>Teacher Training</td>
<td>School Outcome</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Halfors (2006)</td>
<td>Substance use</td>
<td>Randomly selected high risk students from four schools within one school district</td>
<td>695 program, 675 comparison</td>
<td>Teachers indicated, classroom-based course focused on self-esteem, decision making, personal control, and interpersonal communication</td>
<td>School belonging</td>
<td></td>
</tr>
<tr>
<td>Hawkins (1999)</td>
<td>Health-risk and substance use</td>
<td>Individual children randomly assigned to condition classrooms</td>
<td>149 program, 206 comparison</td>
<td>School personnel, universal, school-based teacher training to manage classroom, parent training, and child social skills training</td>
<td>School bonding, school success/failure, school misbehavior</td>
<td></td>
</tr>
<tr>
<td>Horn (2010)</td>
<td>Internalizing symptoms</td>
<td>Classrooms within 6 schools were randomly assigned to condition</td>
<td>201 program, 148 comparison</td>
<td>Research team, universal, school-based CBT and expressive writing activities as a coping strategy</td>
<td>Absences, self-reported grades</td>
<td></td>
</tr>
<tr>
<td>Jaycox (1994)</td>
<td>Depression</td>
<td>Schools non-randomly pre-selected to condition. Equivalence testing of matched participants</td>
<td>69 program, 75 comparison</td>
<td>Research team, indicated, school-based group psychotherapy - cognitive and social problem-solving</td>
<td>Classroom behavior</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>Participants</td>
<td>Condition Details</td>
<td>Staff Details</td>
<td>Outcomes</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------</td>
<td>----------------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>King (1990)</td>
<td>Externalizing symptoms</td>
<td>Children randomly assigned to condition within two schools</td>
<td>30 program, 20 comparison</td>
<td>Paraprofessional Indicated, community-based social skills groups for children, teacher consultation weekly, parent meetings by appointment</td>
<td>Teacher-rated school problems (learning and on-task behavior)</td>
<td></td>
</tr>
<tr>
<td>Kiselica (1994)</td>
<td>Stress inoculation</td>
<td>Children randomly assigned to condition</td>
<td>24 program, 24 comparison</td>
<td>Mixed-School counselor and research team</td>
<td>Universal, school-based course teaching progressive muscle relaxation, cognitive restructuring, and assertiveness</td>
<td>GPA</td>
</tr>
<tr>
<td>Klein (2004)</td>
<td>Internalizing symptoms</td>
<td>Classrooms randomly assigned to condition within three schools</td>
<td>34 program, 28 comparison</td>
<td>Researcher</td>
<td>Selected, school-based group curriculum focused on learning and practicing stress management techniques associated with academic and social stress of gifted children</td>
<td>Academic stress</td>
</tr>
<tr>
<td>Study</td>
<td>Focus</td>
<td>Assignment Type</td>
<td>N</td>
<td>Intervention</td>
<td>Outcomes</td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>------------------------------</td>
<td>---------------------------</td>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td></td>
</tr>
<tr>
<td>Lang (2009)</td>
<td>Externalizing symptoms</td>
<td>Randomly assigned to condition</td>
<td>25 program, 21 comparison</td>
<td>Research team Universal, community-based program teaching psychosocial skills through technological activities</td>
<td>Academic motivation, academic value</td>
<td></td>
</tr>
<tr>
<td>Lochman (2009)</td>
<td>Antisocial behaviors (substance use and aggression)</td>
<td>Children randomly assigned to condition</td>
<td>61 program, 63 comparison</td>
<td>Research team Universal and indicated, school-based intervention providing parent training, child social-cog skills, and teacher consultation</td>
<td>Academic competence, school bonding</td>
<td></td>
</tr>
<tr>
<td>Metropolitan Area Child Study Research Group (2002)</td>
<td>Aggression</td>
<td>16 school randomly assigned to conditions, high-risk youth assigned to conditions non-randomly</td>
<td>1282 program 1085 comparison</td>
<td>Mixed - Research team and community personnel</td>
<td>Selected and indicated intervention: Level A - Teacher training classroom management/Teacher-led social skills intervention to children; Level B - A components + small group peer prosocial training; Level C - A and B components + small group family intervention skills/communication and support network</td>
<td>Achievement tests</td>
</tr>
<tr>
<td>Murray (2005)</td>
<td>Internalizing and externalizing symptoms</td>
<td>Students randomized to condition after being nominated</td>
<td>24 treatment, 24 control</td>
<td>Teachers</td>
<td>Indicated, weekly teacher-student meetings, increased teacher praise, teacher-parent monthly phone calls</td>
<td>GPA, classroom engagement, attendance</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>-------------------------</td>
<td>----------</td>
<td>---------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Neace (2012)</td>
<td>Violence/Aggression</td>
<td>Pre-post matched non-equivalent control group design matched by school and by individual students within schools.</td>
<td>2147 program, 2118 comparison</td>
<td>Teachers</td>
<td>Universal, school-based curriculum teaching empathy, impulse control, emotion/anger management,</td>
<td>Attendance, suspensions</td>
</tr>
<tr>
<td>Shapiro (2002)</td>
<td>Violence/Aggression</td>
<td>No randomization, condition assigned at the school level</td>
<td>1015 program, 415 comparison</td>
<td>Teachers</td>
<td>Universal, school-based curriculum psycho-education on violence attitudes, values, self-concept, and anger management</td>
<td>Suspensions, disciplinary referrals</td>
</tr>
<tr>
<td>Sorrenti (1996)</td>
<td>Externalizing symptoms</td>
<td>Children randomly assigned to condition</td>
<td>11 program, 10 comparison</td>
<td>Researcher</td>
<td>Indicated, school-based groups focused on pro-social behavior and attribution training</td>
<td>Discipline, grades, school attitudes</td>
</tr>
<tr>
<td>Study</td>
<td>Type of Symptoms</td>
<td>Schools Assignment</td>
<td>Program Size</td>
<td>Intervention Description</td>
<td>Outcomes</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------</td>
<td>--------------------</td>
<td>--------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>Stolberg (1994)</td>
<td>Internalizing and externalizing symptoms</td>
<td>Randomly assigned to condition</td>
<td>28 program, 22 comparison</td>
<td>Mixed - School faculty and research team</td>
<td>School attitudes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Selected, school-based groups: 1) support group only - peer support group; 2) support + skill building - labeling feelings; skills transfer - parent workshops and support; 3) support + skills +transfer - parent workshops and workbook for parent-child interaction improvement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storer (1994)</td>
<td>AD/HD symptoms</td>
<td>Experimental pre-post design. Children randomly assigned</td>
<td>8 program, 7 comparison</td>
<td>Researcher</td>
<td>Indicated, school-based group CBT intervention targeting impulsive behaviors</td>
<td>academic self-esteem</td>
</tr>
<tr>
<td>Study</td>
<td>Description</td>
<td>Sample Size</td>
<td>Treatment</td>
<td>Outcome Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stormshak (2009)</td>
<td>Promote well-being and prevent externalizing behaviors and substance use</td>
<td>Youth randomly assigned at the individual level to condition</td>
<td>500 program, 498 comparison</td>
<td>Clinical therapist, Universal and indicated, family resource center to promote positive parenting practices, selected/indicated intervention also available to all - brief intervention based on motivational interviewing to improve parent-child interactions</td>
<td>GPA, Absence</td>
<td></td>
</tr>
<tr>
<td>Suter (1989)</td>
<td>&quot;School adjustment problems&quot;</td>
<td>Random assignment at individual child</td>
<td>14 program, 12 comparison</td>
<td>Teachers</td>
<td>Indicated, teacher-child mentoring program involving coordinated play activities</td>
<td>Achievement tests, school competence subjective</td>
</tr>
<tr>
<td>Timmons-Mitchell (2006)</td>
<td>Antisocial behavior</td>
<td>randomly assigned adolescents to condition</td>
<td>48 program, 45 comparison</td>
<td>Mental health professionals</td>
<td>Indicated, clinic-based multi-systemic therapy- family focused intense contact with family to help empower family to manage child's behavior</td>
<td>School Attitudes</td>
</tr>
<tr>
<td>Study (Year)</td>
<td>Intervention Type</td>
<td>Outcome Measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Analysis of Academic Outcome Categories

Effect sizes were computed for school outcome categories as well. Outcome category formation was guided by outcome categories examined in Baskin et al. (2010) and Becker et al. (2014) research syntheses of mental health programs with school outcomes and by available data from included studies. Four outcome categories were subsequently generated that spanned objective (i.e., school records) and subjective reports (e.g., teacher- parent- and student reports). These included: performance (e.g., GPA, standardized test scores, subject-specific grades), attendance (e.g., days absent/present, tardiness), discipline (e.g., suspensions, disciplinary referrals, etc.), teacher-reported class behavior (e.g., on-task behavior), and child-, parent-, and teacher-reported school attitudes (e.g., school connectedness, academic motivation). All outcome effect sizes were significantly different from zero, except the discipline outcome category, which was marginally significant (p = .055). Outcome categories did not significantly differ from one another. Thus, there was no evidence to support Hypothesis Ib.; that is proximal outcomes did not differ significantly from more distal outcomes. Outcome category effect sizes, corresponding 95% confidence intervals, and p-values are listed in Table 3.
Table 3

*Effect sizes for outcome categories*

<table>
<thead>
<tr>
<th>Outcome category</th>
<th>N of independent samples</th>
<th>Effect size (Hedge’s g)</th>
<th>95% Confidence Interval</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>29</td>
<td>.242</td>
<td>.121 to .362</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Attendance</td>
<td>16</td>
<td>.179</td>
<td>.073 to 285</td>
<td>.001</td>
</tr>
<tr>
<td>Discipline</td>
<td>11</td>
<td>.214</td>
<td>-.005 to 433</td>
<td>.055</td>
</tr>
<tr>
<td>Class Behavior</td>
<td>12</td>
<td>.157</td>
<td>.015 to 300</td>
<td>.031</td>
</tr>
<tr>
<td>School Attitudes</td>
<td>15</td>
<td>.223</td>
<td>.103 to .334</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>
Moderators of Program Effects

To answer Research Question 2, moderator analyses were conducted to determine factors that influence program effects. First, to determine whether moderation analysis is warranted, heterogeneity among samples must exist. The Q-statistic and corresponding p-value were used to test the null hypothesis that all variance among samples is due to random error and is not due to real differences in sample effects (Borenstein et al., 2009). A significant Q-statistic indicates that the studies are not from a common population, while a non-significant Q value indicates the opposite (Higgins, Thompson, Deeks, & Altman, 2003). In the current meta-analysis, the aggregate effect size across 46 samples of all school outcomes was not internally homogenous, $Q(44) = 293.159$, $p < .001)$. Thus, the null hypothesis is rejected, suggesting significant heterogeneity among studies, or that the difference across independent samples is due to real differences in sample effects. As a complement to the Q-statistic, the $I^2$ statistic indicates the percent of heterogeneity among a set of studies (Borenstein et al., 2009). The $I^2$ values range from 0% to 100%. According to Higgins and colleagues (2003), values around 15% reflect a mild degree of heterogeneity, between 25% and 50% a moderate degree, and values greater than or equal to 75% a high degree of heterogeneity. The $I^2$ among the 46 samples included in this meta-analysis is 84.991, indicating that approximately 85% of the variance is due to real sample effects (not random error), and therefore, moderator analysis could explain up to 85% of sample heterogeneity. Based on the significant Q-statistic and high $P$ value, moderator analyses were justified. Borenstein and colleagues (2009) noted that power to
detect the relationship between subgroup membership and effect size or between covariate values and effect size is commonly low. Conclusions drawn from the following moderator analyses should, therefore, be made with caution.

**Moderation with Categorical Variables**

Moderator analyses with categorical moderator variables were conducted to compare effect sizes between groups of studies. More specifically, a mixed effects analysis was used. In a mixed effects analysis, a random effects model is used to combine samples within each group, and a fixed effect model is used to combine groups and yield the overall effect. The sample-to-sample variance (tau-squared) is assumed to be the same for both/all groups; this value is computed within groups and then pooled across groups (i.e., obtaining a pooled variance) (Borenstein et al., 2009). In mixed effects analysis, differences between groups of samples (i.e., moderation) were examined by computing a Q-statistic and corresponding p-value. Analogous to a standard ANOVA, here, the Q-statistic is a test of the null hypothesis that there is no difference between groups.

The first moderator analysis compared studies which were universal in nature to those which had selected participants based on elevated risk. Programs that were selected prevention (i.e., targeting vulnerable youth, but not formally screened for elevated risk) were subsumed under the category of elevated risk due to the small number of studies ($n = 3$) that were selective in nature. Twenty-one samples targeting all youth and 25 samples targeting youth at heightened risk were included in this analysis. Using a mixed effects estimate, the 21 universal program samples resulted in a Hedge’s $g$ and a corresponding 95% confidence
interval of .152 (.038 to .266), and the 25 indicated and selected prevention samples resulted in a Hedge’s g of .260 (.136 to .385). Moderator analysis yielded, \(Q(1) = 1.572, p = .210\), indicating that there was no significant difference between universal and targeted prevention programs.

Table 3 presents findings for all categorical moderator variables, all following the above methodology. In cases in which data were missing for the particular moderator of interest, these studies were excluded from the analyses, resulting in a total independent sample size less than 46. As seen in Table 4, none of the categorical moderator variables (i.e., prevention level, multi-systemic intervention, program setting, and SES) explained study variation significantly.
Table 4

*Results of moderator analyses with categorical moderators*

<table>
<thead>
<tr>
<th>Moderator</th>
<th>Category</th>
<th>N of Studies</th>
<th>Effect Size (g)</th>
<th>95% CI</th>
<th>Q, p Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal vs. Indicated</td>
<td>Universal</td>
<td>21</td>
<td>.152</td>
<td>.038 to .266</td>
<td>1.572, p = .210</td>
</tr>
<tr>
<td></td>
<td>Indicated</td>
<td>25</td>
<td>.260</td>
<td>.136 to .385</td>
<td></td>
</tr>
<tr>
<td>Multi-systemic vs. Individual</td>
<td>Multi</td>
<td>25</td>
<td>.241</td>
<td>.144 to .339</td>
<td>1.884, p = .170</td>
</tr>
<tr>
<td></td>
<td>Individual</td>
<td>21</td>
<td>.129</td>
<td>.002 to .256</td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>Yes</td>
<td>41</td>
<td>.188</td>
<td>.106 to .270</td>
<td>1.080, p = .299</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>5</td>
<td>.343</td>
<td>.062 to .625</td>
<td></td>
</tr>
<tr>
<td>60% low-income</td>
<td>Yes</td>
<td>16</td>
<td>.144</td>
<td>.017 to .271</td>
<td>.893, p = .345</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>11</td>
<td>.236</td>
<td>.095 to .377</td>
<td></td>
</tr>
</tbody>
</table>

**Moderation with Continuous Moderator Variables**

A regression-based analysis, called *meta-regression*, was used to estimate the impact of continuous study moderators (i.e., program length and age of participants) on overall heterogeneity. Meta-regression examines the influence of covariates (moderators) on outcome effects (i.e., effect sizes). A mixed effects
approach was assumed using an unrestricted maximum likelihood model. Again, a mixed effects model, rather than a fixed effects model, allows for within and between study variation; therefore, it is the most appropriate model for these analyses.

Analogous to a standard regression, meta-regression produces and examines a regression line: \( y = a + bx \), where \( x \) is the covariate (moderator) under consideration, \( y \) is the regressed outcome (effect size), \( a \) is the intercept (the effect size when the value of the moderator equals zero), and \( b \) is the slope of the line. If the slope \( b \) is significantly greater than zero, the moderator is said to have a significant effect on the outcome. The statistical program CMA suggests the use of one outcome variable per study as the dependent variable and cannot compute the regression coefficient using the aggregate of multiple effect sizes within study. Thus, given that the academic performance variable was the most commonly reported outcome variable in this study (\( n = 29 \)) and most commonly used variable in the meta-analytic literature of mental health programs and school outcomes, moderators were used in a regression model predicting performance effect sizes.

Initially, program dosage and participant age were proposed as continuous moderators. Analyses were conducted for program length, rather than dosage because very few studies reported any measure of program dosage and across studies there was inconsistency in how programs reported dosage. Participant age was excluded from meta-regression analyses because of significant missing data for child and adolescent participants’ ages. Fifty-four percent of the samples did
not report youths’ age, leaving just 16 studies and 18 independent samples with age data. This low sample size coupled with the focus on performance only outcomes reduced the total sample for this analysis to 12 studies. Thus, meta-regression results examining participant age on program effects would not be representative.

In regards to program length, PMH programs ranged in length from 4 weeks to 312 weeks (M = 43.91, SD = 61.24). Results showed no significant moderation of program length on overall program effect size (b = .001, SE = .001, Z = 1.151, p = .250).

**Supplemental Analyses**

Additional analyses were conducted to examine whether other potential coded variables moderated program effects. These variables were: 1) inclusion of a parent intervention component; 2) mode of delivery (i.e., programs that were primarily delivered by the researcher or research team and programs that were delivered by teachers or other community personnel); 3) over 60% ethnic minority sample (i.e., programs with 60% or greater ethnic minority sample versus those with 60% or greater White sample). Studies that reported including a component that intervened at the level of the parent showed a trend towards significantly moderating program effects, such that programs with a parent component yielded better school outcomes compared to those that did not include parents. All other categorical variables were not found to significantly moderate program effects.
Table 5

Results of supplemental categorical moderators

<table>
<thead>
<tr>
<th>Moderator</th>
<th>Category</th>
<th>N of Studies</th>
<th>Effect Size (g)</th>
<th>95% CI</th>
<th>Q, p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent Component</td>
<td>Yes</td>
<td>14</td>
<td>.280</td>
<td>.159 to .400</td>
<td>2.715, p = .09</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>32</td>
<td>.154</td>
<td>.067 to .240</td>
<td></td>
</tr>
<tr>
<td>Researcher</td>
<td>Yes</td>
<td>12</td>
<td>.147</td>
<td>-.034 to .327</td>
<td>.266, p = .606</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>26</td>
<td>.201</td>
<td>.102 to .299</td>
<td></td>
</tr>
<tr>
<td>60% or more racially</td>
<td>Yes</td>
<td>28</td>
<td>.186</td>
<td>.091 to .280</td>
<td>.731, p = .392</td>
</tr>
<tr>
<td>diverse</td>
<td>No</td>
<td>15</td>
<td>.266</td>
<td>.108 to .423</td>
<td></td>
</tr>
</tbody>
</table>

Publication Bias

Publication bias refers to a phenomenon common in meta-analyses that occurs when research findings in the published literature are systematically unrepresentative of the total population of completed studies (Borenstein et al., 2009). When publication bias exists, conclusions drawn from the published literature may be inaccurate, specifically, an overestimate of the true effect. One hypothesized reason for publication bias is the “File Drawer Effect” (Rosenthal, 1979). This theory posits that statistically significant results are more likely to be
published than null findings, thus biasing the literature base and, consequently, meta-analyses. Another potential reason for publication bias is the tendency for smaller studies to be conducted more rigorously and with greater methodological control (Borenstein et al., 2009). In the current meta-analysis, the 46 included samples were tested for whether they represented a biased sample of all studies. The following statistical procedures were conducted to analyze the potential for publication bias: forest plot, funnel plot, rank correlation, regression, fail-safe $N$, and the trim and fill method.

**Forest Plot**

The forest plot presents a visual representation of the relative weights associated with each independent sample (Borenstein et al., 2009). The plot presents samples with the lowest weight contribution (i.e., smallest sample sizes and largest standard errors) at the top. As seen in Figure 1, there is some evidence to suggest that studies with smaller samples sizes, thus smaller weights, have greater effect sizes than the studies with larger weights and larger sample sizes. This may indicate publication bias.
Figure 1

Effect Size Forest Plot Across 46 Independent Samples

<table>
<thead>
<tr>
<th>Study name</th>
<th>Subgroup within study Outcome</th>
<th>Time point</th>
<th>Hedge's g and 95% CI</th>
<th>Relative weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hâre (1984)</td>
<td>Low Emo Anxiety</td>
<td>Combined</td>
<td>post</td>
<td>0.37</td>
</tr>
<tr>
<td>Hâre (1984)</td>
<td>High Emo Anxiety</td>
<td>Combined</td>
<td>post</td>
<td>0.31</td>
</tr>
<tr>
<td>Sorer (1964)</td>
<td>Blank</td>
<td>School Attitudes</td>
<td>post</td>
<td>0.58</td>
</tr>
<tr>
<td>Castro-Villanueva (2003)</td>
<td>girls</td>
<td>Attendance</td>
<td>post</td>
<td>0.36</td>
</tr>
<tr>
<td>King (1990)</td>
<td>3-4</td>
<td>Combined</td>
<td>post</td>
<td>0.71</td>
</tr>
<tr>
<td>Garrott (1990)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>0.74</td>
</tr>
<tr>
<td>King (1990)</td>
<td>4-2</td>
<td>Combined</td>
<td>post</td>
<td>0.73</td>
</tr>
<tr>
<td>Sykes (1990)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>0.86</td>
</tr>
<tr>
<td>Bursi (1980)</td>
<td>girls</td>
<td>Combined</td>
<td>post</td>
<td>0.96</td>
</tr>
<tr>
<td>Castro-Villanueva (2003)</td>
<td>boys</td>
<td>Attendance</td>
<td>post</td>
<td>1.06</td>
</tr>
<tr>
<td>Archuleta (1988)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>1.15</td>
</tr>
<tr>
<td>Long (2005)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>1.35</td>
</tr>
<tr>
<td>Murray (2005)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>1.31</td>
</tr>
<tr>
<td>Kielczak (1994)</td>
<td>Blank</td>
<td>Performance</td>
<td>post</td>
<td>1.31</td>
</tr>
<tr>
<td>Stoddard (1994)</td>
<td>trans skills &amp; sup</td>
<td>Performance</td>
<td>post</td>
<td>1.33</td>
</tr>
<tr>
<td>Bursi (1980)</td>
<td>boys</td>
<td>Combined</td>
<td>post</td>
<td>1.35</td>
</tr>
<tr>
<td>Klein (2004)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>1.50</td>
</tr>
<tr>
<td>Jaycox (1986)</td>
<td>Blank</td>
<td>Class Behavior</td>
<td>post</td>
<td>1.53</td>
</tr>
<tr>
<td>Tramino-Mitchell (2006)</td>
<td>Blank</td>
<td>School Attitudes</td>
<td>post</td>
<td>1.78</td>
</tr>
<tr>
<td>Letchman (2002)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>2.14</td>
</tr>
<tr>
<td>Cappella (2012)</td>
<td>blank</td>
<td>School Attitudes</td>
<td>post</td>
<td>2.27</td>
</tr>
<tr>
<td>Gottfredson (2002)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>2.44</td>
</tr>
<tr>
<td>Egger (1989)</td>
<td>blank</td>
<td>Combined</td>
<td>post</td>
<td>2.47</td>
</tr>
<tr>
<td>Walker (2003)</td>
<td>blank</td>
<td>Combined</td>
<td>post</td>
<td>2.54</td>
</tr>
<tr>
<td>August (2003)</td>
<td>Blank</td>
<td>Performance1</td>
<td>post</td>
<td>2.55</td>
</tr>
<tr>
<td>MACS (2002)</td>
<td>Level B Late</td>
<td>Performance</td>
<td>post</td>
<td>2.54</td>
</tr>
<tr>
<td>Breaston (2012)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>2.88</td>
</tr>
<tr>
<td>MACS (2003)</td>
<td>Level C Late</td>
<td>Performance</td>
<td>post</td>
<td>2.70</td>
</tr>
<tr>
<td>MACS (2003)</td>
<td>Level A Late</td>
<td>Performance</td>
<td>post</td>
<td>2.74</td>
</tr>
<tr>
<td>Hawkins (1999)</td>
<td>full v control</td>
<td>Combined</td>
<td>post</td>
<td>2.92</td>
</tr>
<tr>
<td>Tall (2004)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>2.97</td>
</tr>
<tr>
<td>Stormshak (2009)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>2.97</td>
</tr>
<tr>
<td>Cataldo (2003)</td>
<td>girls</td>
<td>Combined</td>
<td>post</td>
<td>3.06</td>
</tr>
<tr>
<td>MACS (2003)</td>
<td>Level A Early</td>
<td>Performance</td>
<td>post</td>
<td>3.15</td>
</tr>
<tr>
<td>MACS (2003)</td>
<td>Level B Early</td>
<td>Performance</td>
<td>post</td>
<td>3.15</td>
</tr>
<tr>
<td>Coleman (2000)</td>
<td>Blank</td>
<td>Discipline</td>
<td>post</td>
<td>3.21</td>
</tr>
<tr>
<td>Shapiro (2002)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>3.42</td>
</tr>
<tr>
<td>Fay (2001)</td>
<td>Nevada v matched</td>
<td>Combined</td>
<td>post</td>
<td>3.46</td>
</tr>
<tr>
<td>Nescio (2012)</td>
<td>Cohort 1</td>
<td>Combined</td>
<td>post</td>
<td>3.51</td>
</tr>
<tr>
<td>Gottfredson (1993)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>3.50</td>
</tr>
<tr>
<td>Haffner (2006)</td>
<td>Blank</td>
<td>Combined</td>
<td>post</td>
<td>3.53</td>
</tr>
</tbody>
</table>
Funnel Plot

The funnel plot is a plot of the measure of sample standard error on the vertical axis as a function of Hedge’s g on the horizontal axis. When samples are distributed symmetrically about the combined effect size, publication bias is absent. When the bottom of the plot shows a higher concentration of samples on one side of the mean than on the other, publication bias is present (Borenstein et al., 2009). In the current meta-analysis, there is some dispersion of samples at the bottom toward the right-hand side of the graph, suggesting the possibility of publication bias (see Figure 2).

Figure 2

*Funnel plot of standard error by standard difference in means*

Begg and Mazumdar Rank Correlation Test

To capture the bias represented by the funnel plot mentioned above quantitatively, Begg and Mazumdar (1994) suggested that this inverse correlation
between standard error (sample size) and effect size can be computed and serve as a test of publication bias. Specifically, a rank order correlation (Kendall’s tau b) between the treatment effect and the standard error is computed. A significant correlation suggests the existence of bias. In the current analysis, Kendall’s tau \( b = .039, Z = .381, p(1\text{-tailed}) = .351, p(2\text{-tailed}) = .351; \) therefore, the rank correlation test does not indicate significant publication bias.

**Egger’s Regression Test**

Similarly, Egger’s linear regression method (Egger, Davey Smith, Schneider, & Minder, 1997) is also intended to quantify the bias captured by the funnel plot. Egger, however, suggests using the actual values of the effect sizes and their precision, rather than ranks, by regressing the standardized effect on the inverse of the standard error. In the resulting regression equation, the slope represents the treatment effect, and the intercept is a measure of bias. A significant intercept suggests the existence of bias. In the current analysis, Intercept = 1.131, \( SE = .561, CI95 = -.001 \) to 2.263, \( t(43) = 2.01, p(1\text{-tailed}) = .025, p(2\text{-tailed}) = .050. \) These p-values suggest marginally significant to significant publication bias.

**Fail-Safe N**

If publication bias is present, it is hypothesized that some non-significant studies are missing from our analysis, and including these missing studies would nullify the observed effect. Therefore, the number of studies that would be required to nullify the effect – the Fail-safe \( N \) (FSN) – is computed. As reported in the above results, this meta-analysis incorporates data from 45 studies, which
yield a z-value of 10.465 and corresponding p-value less than 0.001. The FSN is 1,238, which means that 1,238 null studies (mean Hedge’s g = 0) would need to be located and included in order for the combined p-value to exceed 0.05. More conservatively estimated, when the alpha level was set to 0.01 (instead of 0.05), analysis yielded a FSN of 698.

Rosenthal (1979) suggested that the FSN be equal to or larger than five times the number of retrieved studies (or, in this case, independent samples) plus 10. Both FSN estimates in this meta-analysis exceed Rosenthal’s recommended resistance number, 45 x 5 + 10 = 235, thus indicating no significant bias.

**Duval and Tweedie’s Trim and Fill**

Based on the four methods above, there is some evidence of publication bias. Next, it is important to ask how the intervention effect (overall effect size) would shift if bias were to be removed. In reference to the funnel plot, because a relatively high number of small samples (with large effect sizes) fall toward the right of the mean and relatively few fall toward the left, there is concern that these “left-hand” studies may actually exist and are missing from the analysis. Duval and Tweedie (2000) developed a method that allows for the imputation of these studies, called Trim and Fill. That is, the theoretical locations of these missing studies are determined, the studies are added to the analysis, and then the combined effect is recomputed. In the current analysis, assuming a random effects model of imputation, the trim and fill method suggested that no studies are needed to remove bias. Thus Hedge’s g and corresponding 95% confidence interval remains this same as previously stated 0.202 (.121 to .283).
In sum, examining the forest and funnel plots visually appears to suggest some potential for publication bias based on the mixed findings from the rank correlation and the intercept tests. However, even the most conservative estimate of the fail-safe N suggests that 698 studies with null findings would need to be found in order to bring the overall effect size to a non-significant level. Illustrated proportionally, for every one of the 46 observed samples in this meta-analysis there would need to be 15 missing null samples for the overall effect to be nullified. Further adding to the evidence of minimal publication bias in the current meta-analysis, the trim and fill method indicates that no studies need to be added to remove bias. Taken together, findings in this meta-analysis appear to be fairly robust.
CHAPTER IV
DISCUSSION

This meta-analysis sought to explore the potential for youth PMH programs to improve school outcomes. Forty-six independent samples across 34 studies were identified in the literature as meeting criteria for inclusion in the current meta-analysis. Of note, most studies were excluded simply because they did not report at least one explicit school outcome (e.g., grades, attendance, school-related attitudes). The tendency for many PMH intervention studies to exclude even a cursory measurement related to school outcomes represents a significant oversight in the current state of the field. The implications of this oversight will be discussed later in this section.

An examination of key study characteristics revealed that most PMH programs were designed in a manner consistent with preventive mental health theory (Bronfenbrenner, 1994; Coie et al., 1993; Greenberg, Domitrovich, Bumbarger, 2000). Most PMH programs intervened at the microsystemic level, for example, within the school setting (85.7%). Other programs took a mesosystemic approach by linking microsystems (e.g., school-family interventions). Across ecological systems, PMH programs reported intervening in efforts to reduce multiple risk and protective factors for youth from ethnically and economically diverse backgrounds. The variability across PMH interventions is consistent with the multi-pronged approaches to prevention of youth mental disorders as described by Greenberg and colleagues (2000). A more thorough discussion of the heterogeneity of studies is discussed later in this section.
Results across 46 independent samples of universal, selected, and indicated PMH interventions yielded a small but significant effect size (Hedge’s $g = .202$), providing evidence that PMH programs for youth can yield benefits relevant to educational outcomes. Further analysis of individual school outcome categories (i.e., performance, attendance, discipline, class behavior, and school attitudes) demonstrated that effect sizes remained significantly different from zero (or marginally significant for the discipline category) and all were within a similar effect size range as the overall school outcome effect size. Taken together, these results suggest PMH may provide positive benefits outside of the mental health domain, specifically, improving academic performance, attendance, class behavior, school attitudes, and possibly reduce disciplinary problems in schools.

**Mental Health, Mental Illness, and Positive Youth Development**

Previous meta-analyses of PMH and mixed prevention and treatment programs reported school outcome effect sizes between 0.0 to .43 (e.g., Prout & Prout, 1998; Durlak et al., 2011; Baskin et al., 2010; Payton et al., 2008). The small effect size of the current study (.201) is consistent with Durlak et al. (2011) and Baskin et al. (2010) whom reported overall small effect sizes for school outcomes of .28 and .38, respectively. This meta-analysis adds to the general trend for school outcome effect sizes to be generally smaller than previous meta-analyses’ mental health outcome effect sizes, which are generally in the large range (e.g., Durlak & Wells, 1997, 1998). It is likely that the discrepancy between school and mental health outcome effect sizes from past syntheses is attributable to the fact that school improvement was not the direct target of any of
these mental health interventions. Thus, it would be expected that the target of the intervention would yield greater effects than an indirect target (i.e., school outcomes). However, the variability from null to moderate effect sizes of school outcomes across multiple meta-analyses is curious and warrants further consideration.

One possible explanation for the variability in findings across meta-analyses is the wide range of variability in inclusion/exclusion criteria across studies. As previously mentioned, for example, several meta-analyses limited inclusion to performance-related outcomes only (e.g., Durlak & Wells, 1997; Durlak et al., 2011, Prout & Prout, 1998). Other meta-analyses differed in the focus of the programs included. For instance, Durlak et al. (2011) included school-based programs that emphasized at least one SEL component (e.g., leadership development programs). Finally some meta-analyses included all mental health interventions, that utilized counseling or therapy, regardless of whether youth were diagnosed with a disorder or not (e.g., Prout & Prout, 1998; Baskin et al., 2011), thus these meta-analyses had mixed prevention and treatment samples. The current meta-analysis was broader in some regards (i.e., broad school outcomes, not limited to school-based interventions) and narrow in other respects (i.e., prevention only, explicit target of promoting mental health to prevent a disorder or mental illness symptom reduction). The variability in inclusion criteria across studies brings up theoretical questions regarding the ambiguous nature by which studies are included or excluded in syntheses.
Weare and Nind (2011) conducted a systematic review of meta-analyses and systematic reviews (a review of reviews) focused on mental health promotion programs. In their review, authors took a broad approach to the definition of mental health. Thus, they included meta-analyses in their review that were focused on specific prevention of disorders and symptoms as well as programs that promoted general social competencies, such as Durlak and colleague’s SEL meta-analysis and Catalano and colleague’s (2002) review of positive youth development programs that generally promote character development in youth. Weare and Nind (2011) argued that casting a wide net of inclusion in their review of reviews allowed for a more comprehensive picture of mental health promotion. Future studies should specifically explore the variability in school outcome effect sizes of previous meta-analyses (a meta-analysis of meta-analyses) in an effort to determine whether effect sizes vary in some systematic manner. An interesting empirical question may be to determine whether programs that focus on general character development and specific SEL skills differ significantly from programs that focus more specifically on the prevention of a mental illness and reduction of symptoms.

The variability in definitions and conceptualizations of PMH across meta-analyses and reviews warrants an even broader discussion on what it means to promote mental health versus prevent mental illness. In their seminal meta-analysis of PMH programs, Durlak & Wells (1997, 1998) defined PMH as interventions that seek to promote mental health and/or prevent mental illness. It is commonly understood that mental illness is identified by the presence of
symptoms of psychopathology, as defined by nosological classifications. Mental health, however, is less defined in the current literature, thus presenting numerous challenges in research when one attempts to capture the concept. The World Health Organization (WHO 2005, p. 2) defined mental health as “a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community.” Given this definition, the kinds of interventions that might promote “mental health” may vary widely.

As theories of positive psychology have emerged through the years, research is suggesting that mental health and mental illness may be very distinct and separate concepts, even having different predictive implications (Keyes, 2002, 2004, 2005a, 2005b). Extensive research has been conducted in the past decade to define mental health more clearly. As mentioned previously in this dissertation, strides have been made in positive psychology that promote the idea that one may be mentally healthy (flourishing) or mentally unhealthy (languishing), and that being mentally unhealthy does not necessarily mean meeting criteria for a mental illness (Westerhof & Keyes, 2010). In this model, mental health is defined by (1) well-being, (2) effective functioning at the individual level, and (3) effective function at the community level. This model provides a better fit for SEL programming which specifically seeks to promote social and emotional skills and competencies that promote positive youth development (Zins, 2004) and as a consequence, may serve as protective factors against the development of mental illness. Whereas, many other PMH programs
may be more targeted by focusing on specific risk factors associated with the development of mental illness. For example, Durlak and colleagues (2011) included programs such as leadership development in their meta-analysis of SEL programs, while one of the studies included in the current meta-analysis (Hains, 1994) was a stress inoculation intervention that sought to prevent internalizing symptoms in youth. Comparing between the more broad SEL approach and the more focused stress inoculation example highlights areas of complement and contrast associated with mental health promotion and mental illness prevention. An important charge for the field of PMH would be to become clear and consistent in our conceptualization of mental health and what it means to promote mental health.

**Moderators**

This is the first meta-analysis to date to systematically review and synthesize findings specific to PMH program effects on school outcomes. Indeed, past meta-analyses have examined overall effectiveness of social-emotional learning (i.e., Durlak et al., 2011) or general PMH (i.e., Durlak & Wells, 1998). However, no meta-analysis to date has comprehensively examined program effects as well as moderators of program effectiveness specific to school outcomes.

Proposed categorical and continuous moderators were not found to moderate program outcomes. As stated in the Results Section, moderation analyses generally have low statistical power, thus conclusions should be understood with caution.
**Prevention Level**

No significant difference was found between universal and indicated programs. Previous research has suggested that children and adolescents who are already evidencing a mental health problem are likely to benefit most, that is, evidence the greatest change, compared to a universal population of youth (Durlak & Wells, 1998, 1999). The current study suggests that overall PMH is just as beneficial among universal and indicated samples of youth. Reflecting on the work of Suldo and Shaffer (2008) on “flourishing” and “languishing” youth helps to understand why this may be the case. In their study of youth between 10 to 16 years of age, they found poor academic performance was associated with 17% of their sample who were symptomatic and had low subjective well-being, as well as 13% of their sample who were not symptomatic but had low-subjective well-being. The current meta-analysis lends some evidence to suggest that screening procedures to detect mental illness symptoms may miss youth who would also benefit from PMH programming, despite not having symptoms.

**Program Characteristics**

No differences were found between programs that took a multi-systemic approach (i.e., focused change on an element of the child’s environment) versus those that focused just on the individual child. The non-significance of this moderator is intriguing because it suggests that simply effecting change at different environmental systems is not inherently better than just focusing on the individual child. Two additional hypotheses can be generated as well: 1) effects associated with multi-systemic components vary by the target (e.g., parents,
teachers, policies); 2) important confounds may exist associated with intervention approaches used in multi-systemic programs (e.g., parent-training versus parent support group). The trend towards significance ($p = .09$) suggesting programs that included a parent component had greater positive effects than those that did not include parents, lends support towards Hypothesis 1.

The importance of parental involvement on youth mental health and educational outcomes has been noted throughout mental health and educational literatures (e.g., Patterson, Dishion, & Bank, 1984; Fan & Chen, 2001; Patrikakou, 2008). The trend towards significance of this moderator suggests that involving parents in PMH interventions may also help promote school outcomes. In this investigation, programs that included parents generally sought to enhance parent-child interactions in the format of skill-building (e.g., communication skills), parent support groups, and parent-teacher consultations. Despite the potential benefit of involving parents in interventions, many educators and mental health professionals encounter challenges in successfully engaging parents in intervention efforts. In their chapter focused on the promotion of SEL skills among youth, Patrikakou and Weissberg (2007) reported themes that reflected SEL programs that successfully bridged the gap between schools and families. These themes emphasized (1) the open dialogue between the school and family about program (i.e., SEL) goals to promote buy-in, (2) parent involvement at home through the provision of home-based materials or activities, and (3) parent involvement at school through flexible scheduling and creative school-based activities. Thinking critically about how to partner with parents and engage them
in PMH interventions in the community is a critical element of successful youth programming.

**Setting**

Setting was proposed as a moderator, however, due to the few PMH studies conducted outside the school setting ($n = 5$), the results related to this moderator should be considered with extreme caution. The lack of PMH programs implemented outside of the school setting is likely reflective of the convenience associated with access to the population in research as well as the restrictions that insurance places on conducting preventive interventions in clinical settings (Weist et al., 2014). No differences were found between school-based and non-school-based programs. These analyses should be replicated with a larger sample size to determine whether differences in effects would be statistically significant with the inclusion of more non-school-based studies. Without more studies to constitute this group, further interpretation is ill-advised.

**Child Characteristics**

No significant differences were found between youth samples from low-income backgrounds and those who were not low-income. Nor were differences found between youth samples that were primarily from ethnic minority backgrounds and those samples that were primarily White. This suggests that the programs were just as beneficial to improving school outcomes for youth from diverse backgrounds.
In a previous meta-analysis of mental health treatment effects on school outcomes, researchers found that studies that included primarily children from diverse backgrounds (i.e., ethnic minority) had larger effect sizes compared to studies with primarily White youth (Baskin et al., 2010). The aims of the current study differ from Baskin et al. (2010) in a number of ways, as mentioned in the Introduction Section. The findings from the current meta-analysis are promising because they suggest that the prevention programs included in this study provided benefits to youth regardless of their background. Prevention theory dictates that interventions should heavily weight the contextual elements and individual characteristics of the target population when developing a preventive intervention (Ellis, 1998; Greenberg et al., 1999). Results suggest that overall the PMH programs included in this meta-analysis were tailored appropriately to yield the benefits to their specific target population.

**Length**

A regression analysis examined the continuous variable of program length. Longer programs did not predict a greater academic performance effect size. Several past meta-analyses of mental health interventions have found non-significant findings for program length (e.g., Farahmand et al., 2011). Past studies have noted the importance of measuring prevention program dosage rather than length of the program (Nation, Crusto, Wandersman, Kumpfer, Seybolt, et al., 2003). Dosage refers to an actual measurement of how much participants were exposed to the intervention. A program may last 6 months but participants only receive the intervention once a week for 1 hour per week resulting in a total
dosage of 24 hours. While another program may last 4 weeks total but be a 2-hour sessions three times per week, resulting in a total dosage equal to a program that lasts six times as long. Unfortunately, due to the infrequency and unreliability of studies reporting dosage, dosage was not used in this study and length was used instead. The lack of consistent and reliable measurement of program dosage across participants in studies highlights an area for improvement in the PMH program evaluation literature.

**Efficacy vs. Effectiveness**

No differences were found between programs that were implemented by the researcher and those that were implemented by community personnel. Past research suggests that implementation by the research team would yield greater effect sizes (Durlak & DuPre, 2008). However, the current findings are promising, suggesting that programs that are implemented by community members are just as beneficial as those conducted by a research team. Several explanations for this null finding are possible. Durlak and DuPre (2008) discussed how program implementation plays a significant role in program outcomes. Durlak and DuPre (2008) provided conceptual models of how programs implemented within communities can be implemented properly with maximum benefits. Some of the key factors included training and on-going consultation of community members who were implementing the intervention. In many of the studies included in this meta-analysis that were implemented by community personnel, researchers noted that training was provided to community staff and many included on-going consultation to implement the intervention. Another
explanation why implementation by the researcher was not a significant moderator for program effects could be because teachers or whole-schools implemented the intervention. In these cases, whatever negative effects that would be associated with low fidelity of the intervention by community members may have been outweighed by the benefits of school personnel implementing the intervention. In many cases, school personnel were youths’ classroom teachers. For example, in the Reconnecting Youth Program (Castro-Villarreal, 2009) youths at-risk for substance use and school drop-out were enrolled in an elective course directed by their teacher. Perhaps this personal connection to a school teacher helped to engage youth academically and contributed to the positive effects in school attendance (Hedges g = .893 girls, .203 boys).

**Heterogeneity of Youth PMH**

One additional explanation for the overall lack of moderation may be the high degree of variability among studies. By definition, heterogeneity in meta-analysis effect sizes refers to the variation in the true effect size (Borenstein, Hedges, Higgins, Rothstein, 2009). Again, moderation analyses seek to capture this variation by hypothesizing that specific moderator variables will produce meaningful subgroups that represent true heterogeneity rather than random error. A possible explanation of the results is that when aggregating highly variable interventions across highly variable outcomes, these moderators were not meaningful above and beyond error. Heterogeneity in meta-analyses is common and generally expected, particularly in studies that assume random effects models (Borenstein et al., 2009). However, it is important to reflect on the implications of
such heterogeneity in this meta-analysis and provide recommendations for addressing this issue in the future.

First, there was variability in the types of measures used to assess school outcomes. Even within categories (e.g., performance, attendance, etc.) there were a variety of ways these outcomes were measured between and within studies. In a recent review of educational outcomes measured in mental health treatment and indicated prevention programs, authors wrote extensively about the heterogeneity in school outcome measures. Across 88 studies in the sample, authors reported 45 different measures identified by authors as the “primary indicator of educational outcomes” (Becker et al., 2014, p. 12). Even further, subscales within measures adds additional variability. For instance, GPA could be dismantled by subject to reveal subject specific (e.g., math, reading) effects or non-significant effects. Finally, Becker and colleagues (2014) noted that performance-related outcomes were more likely to be standardized measures (e.g., cognitive tests, standardized statewide tests, etc.), while all other outcome categories (e.g., academic attitude measures, class behavioral measures) were more varied and less likely to be standardized. The heterogeneity of school measures within the current meta-analysis is likely a strong contributor to the lack of homogeneity in study findings.

Second, there was heterogeneity in program approaches to preventing mental health problems in youth. Durlak and Wells (1997) described youth PMH as programs that focus on “reducing the incidence of future adjustment problems…as well as efforts directed at the promotion of mental health functioning” (pg. 116). As mentioned earlier in the Discussion Section, some
studies have taken the position that PMH is broad in scope because it targets any number of risk factors associated with the future mental illness and targets any number of protective factors associated with developing appropriate competencies that may moderate pathways to mental illness. Just as the variations in operational definitions affect the inclusion or exclusion of specific studies, thus causing variability in effect sizes across meta-analyses, variability in the approaches used to prevent problems in youth may be a leading cause of heterogeneity within this meta-analysis.

**Limitations & Future Directions**

The current study is not without limitations. As mentioned previously, although PMH was found to enhance school outcomes overall, there is little clarity as to what moderates program outcomes. Lack of clarity as to what moderates program outcomes presents a barrier to best practices. Some of the problems related to moderation may be related to small overall sample size, as well as even smaller sample sizes for specific moderators (e.g., age, non-school setting). It is also likely that variables that would indeed moderate program effectiveness were not hypothesized in this meta-analysis. Future analyses should examine additional moderators, such as the various intervention components and practices used by the intervention as well as program implementation fidelity (Durlak & DuPre, 2008). Further, more complex data analytic strategies may be useful to control for confounds in the data.

Another set of limitations are related to the studies themselves. Many programs were compared to a control group who received nothing or treatment as
usual (TAU), which was often unspecified. Comparison groups whom received nothing or an ambiguous TAU introduce the question of whether program effects were simply a consequence of placebo. Although the inclusion/exclusion criteria required that the comparison group not receive a different type of PMH programming, effects would be more trustworthy if the TAU was at least more defined. Another area for improvement in the PMH intervention literature is that very few studies reported any type of follow-up. Lack of follow-up makes it difficult to determine the lasting effects of PMH on school outcomes. Within the educational arena, it would be helpful to firmly establish whether these programs must be on-going or can provide lasting benefits when implementing just for a time.

Lastly, the current meta-analysis did not examine mediators of program effects on school outcomes. An important step forward in the literature would be to meta-analyze the extent to which psychological outcomes and school outcomes co-vary. Specifically, at the meta-analytic level, it would be important to determine whether psychological outcome effects and school outcome effects of PMH intervention are correlated. If these outcomes are indeed correlated, it would be important to determine whether psychological outcome effects mediate the relationship between PMH intervention and school outcome effects. Figure 3 presents a potential mediational model.
Testing the proposed model may be premature, however, in the current state of the literature (i.e., limited, inconsistent, and variable measurement of academic outcomes among studies). In Becker and colleagues’ (2014) synthesis of youth mental health studies and academic outcomes, the authors suggested that due to the lack of academic outcome measurement, mediational models at the aggregate level may be difficult to evaluate. For instance, the authors suggested that perhaps a mediational model may be missing key variables such as academic self-efficacy that may further mediate outcomes. In the current meta-analysis most studies evaluated performance-related variables ($n = 30$), and only 15 studies measured any type of school attitude variable. The literature suggests that attitudes generally precede behavioral changes and may serve as an important mediator to those changes. Thus, important limitations exist in the literature that may hinder the evaluation of a comprehensive meta-mediation model. From the standpoint of the interventionist, it presents little added cost and can yield potential benefits to the field to include at least rudimentary academic measures into PMH evaluation studies.

**Implications for Policy and Practice**
Mental health plays an important role in the academic success of youth (Suldo & Shaffer, 2008). Given the link between PMH programming and positive school outcomes, one policy implication of this meta-analysis is to improve the measurement and reporting of school outcomes in PMH intervention research. Inclusion of basic standardized measures would allow researchers to further explore the role mental health intervention plays in academic success. Exploring these relationships would provide school educators and administrators evidence of the relevance PMH interventions have on school outcomes, especially academic performance outcomes.

Another major implication of this meta-analysis is that schools should invest in promoting the mental health of all youth in order to improve school outcomes. This meta-analysis adds to the growing body of literature that provides strong evidence for the causal relationship between mental health intervention and educational benefits for children and adolescents. As reflected in this review, most PMH programs occurred in school where access to children and delivery systems are built into the structure of the school context. School-based mental health initiatives have proliferated the research and practice arena, particularly for PMH approaches. Schools are often ideal because clinic-based prevention is often not billable. Results from the current investigation push the agenda for schools to consider mental health as not just a health interest but also an educational interest – an interest directly relevant to the goals that schools are charged with achieving.

Schools’ not being reliant on research teams’ direct delivery of services would allow for long-term sustainability of PMH programming in the school
context. This meta-analysis included studies of effectiveness, in which school personnel or others indigenous to the community implemented the intervention. Moderation analyses showed no difference between studies implemented by the research team and those implemented by community members. However, many schools are not resourced to effectively implement school-based mental health initiatives. Resources are often a common barrier to the effective implementation and sustainability of new innovations (Aarons, Hurlburt, Horwitz, 2011; Durlak & DuPre, 2008). Programs that are not able to be implemented properly due to a lack of resources are likely to fail and provide skewed data related to the effectiveness of a given program (Durlak & DuPre, 2008). Several models exist in the literature of sustained school-based mental health initiatives. Many sustainable models of school-based PMH are realized through existing school-based health centers. School-based health centers often are operated in partnership between the school and a community health organization (e.g., hospital or local health department). Approximately 2,000 school-based health centers operate nationwide (National Assembly on School-Based Health Care, 2012).

An innovative approach to implementation and sustainment of PMH programs utilizing existing school-based health centers are university and school partnerships (Ward, Strambler, & Linke, 2013; Weist, Stiegler, Stephan, Cox, & Vaughn, 2010). The School Mental Health (SMH) program in Baltimore is a strong example of sustainable SMH programming through university partnerships. Weist and colleagues (2010) described how they were able to enhance general school health centers and school-based mental health services to
a more comprehensive model including high quality prevention efforts. Specifically, the Excellence in School Mental Health Initiative (ESMHI) sought to bring enhanced SMH to participating schools. Enhanced SMH included, building and enhancing community partnerships with other universities, community foundations, and the school district (Weist et al., 2010). These partnerships supported further efforts to boost clinician time on-site, develop relationships with families to increase buy-in, and provide universal prevention services to all youth in the two schools involved in the project. Beginning in 2007 through 2010, the expanded SMH was able to lead to the implementation of several PMH interventions, one of which is included in this meta-analysis (Coping Power; Lochman et al., 2002). Working in conjunction with the university and community partners, the two schools have been able to sustain expanded SMH. Because the article was focused on processes, Weist and colleagues (2010) did not report results of the interventions. Nevertheless, this model of implementing and sustaining PMH programming is promising.

Although PMH appears to show promise as an approach to improving school outcomes among all youth as well as youth evidencing some mental illness symptoms, these effects were modest. The cost-benefit of implementing and sustaining these PMH efforts has yet to be determined. One way to ameliorate this cost-benefit scenario for already over-burdened and often under-resourced schools is to establish community and school partnerships.
References

* Indicates study was included in this meta-analysis


* Coleman (2000) A controlled evaluation of the effects of classroom coping skills training on children's aggressive and externalizing behaviors*


Contributions to positive psychology (pp. 85–103). New York: Oxford University Press.


from three scientific reviews. Chicago: Collaborative for Academic, Social, and Emotional Learning.


### Appendix A

#### Study Characteristics

<table>
<thead>
<tr>
<th>StudyID</th>
<th>Assigned Study ID#</th>
</tr>
</thead>
<tbody>
<tr>
<td>AuthorYr</td>
<td>String Author name and year</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>What type of publication is this?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>book chapter</td>
</tr>
<tr>
<td>2</td>
<td>journal article of book chapter</td>
</tr>
<tr>
<td>3</td>
<td>thesis or doctoral dissertation</td>
</tr>
<tr>
<td>4</td>
<td>technical report</td>
</tr>
<tr>
<td>5</td>
<td>Other [TypeOthSpec]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>qualdesp</th>
<th>Please describe the focus/primary aim of this preventive intervention.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>compser</th>
<th>Prevention focus:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. Internalizing symptoms</td>
</tr>
<tr>
<td></td>
<td>2. Externalizing symptoms</td>
</tr>
<tr>
<td></td>
<td>3. Socio-emotional competence, self-esteem (EXCLUDE)</td>
</tr>
<tr>
<td></td>
<td>4. Alcohol/substance use and/or abuse</td>
</tr>
<tr>
<td></td>
<td>5. Other:</td>
</tr>
<tr>
<td></td>
<td>6. Combination:</td>
</tr>
<tr>
<td></td>
<td>7 = Unclear/Unspecified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TxDom</th>
<th>Treatment type. Please rank (1 = most dominant) in order of dominance (most central components) the intervention style(s) of the program. Look for explicit mention of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>____ Psychoeducation</td>
</tr>
<tr>
<td></td>
<td>____ Parent skill-building</td>
</tr>
<tr>
<td></td>
<td>____ Psychotherapy, Specify</td>
</tr>
<tr>
<td></td>
<td>____ Teacher skill-building</td>
</tr>
<tr>
<td></td>
<td>____ Academic development, Specify</td>
</tr>
<tr>
<td></td>
<td>____ Other skill-building, Specify</td>
</tr>
<tr>
<td></td>
<td>____ Social/Emotional skills, Specify</td>
</tr>
<tr>
<td></td>
<td>____ School-wide capacity/curricular changes</td>
</tr>
<tr>
<td></td>
<td>____ Mentoring, Specify</td>
</tr>
<tr>
<td></td>
<td>____ Other, Specify</td>
</tr>
</tbody>
</table>


1 = central component
2 = secondary component
3 = tertiary component
4 = N/A

[Prevlvl] What was the prevention level?
1 = Universal – administered to all youth regardless of level of risk. For example, all children at a particular community agency or all children in the 2nd grade.
2 = Selected – administered to all youth who may be more vulnerable, based on circumstance, not on measured risk. For example, all adolescents who are transitioning to high school or all children from single-parent homes
3 = Indicated – administered to youth who were assessed and met study criteria for increased levels of psychopathology, but do not meet full criteria for a particular disorder. For example, aggressive children, children who meet some symptoms of depression are not diagnosed as having a depressive disorder.

[Prevlvl_SorI] If the intervention is selected or indicated, specify how it is so (Choose one)

**Selected**

1 = Transition (e.g., transition to high school)
2 = Specific Academic track (e.g., all children who are in remedial classes, were retained the previous year)
3 = Family circumstance (e.g., divorce, single parent)
4 = Disability (e.g., organic or learning disability)
5 = Experienced a traumatic event

**Indicated**

6 = Elevated problem behavior (e.g., aggression, failing grades). Please specify behavior
7 = Elevated mental illness symptoms. Please specify disorder
8 = Other measured risk (e.g., children with current poor grades)
9 = General Other

[Prevlvl_SorIOther] If Other, please specify ____________________ (STRING VARIABLE)

[IntvSet1] Intervention setting [Where did the intervention take place?]

1 = Clinical (e.g., hospitals, clinics)
2 = Community agency
3 = Schools
4 = Home
5 = Residential Treatment Facility
6 = Other (Please specify)
Intervention Setting

1 = Rural  
2 = Urban  
3 = Suburban  
4 = Unknown/unspecified

Please specify the city/state/country if specified

(RSTRING VARIABLE)

Research Design

Unit of assignment to conditions. Select the code that best describes the unit of assignment to treatment and control groups.
1 = Individual (child, teacher, parent, other)  
2 = Classroom, facility  
3 = program area, regions  
4 = cannot tell  
5 = School  
6 = Other

Type of assignment to conditions. Select the code that best describes how subjects were assigned to treatment and control groups.
1 = random after matching, stratification, blocking  
2 = random simple  
3 = nonrandom, post hoc matching  
4 = nonrandom, other (e.g., general comparison group not matched on key characteristics)  
5 = other (specify)

Overall confidence of judgment on how subjects were assigned.
1 = very low  
2 = low (guess)  
3 = moderate (weak inference)  
4 = high  
5 = very high (explicitly stated)

Was the equivalence of the groups tested at pretest?
1 = Yes  
2 = No

Who is the agent of change? Although the treatment group is always the child, who(m) is (are) actually receiving the intervention directly? Select all that apply.
ServD1 Child – Yes = 1, No = 2  
ServD2 Parent/Caregiver – Yes = 1, No = 2  
ServD3 Teacher(s) – Yes = 1, No = 2  
ServD4 Other – Yes = 1, No = 2
ServD4spec _____ Other specify________________________
____( STRING VARIABLE)

Type of Intervention [please check either person centered, environmental focus or mix – if environmental or mix, specify further]

**Person centered:** Work directly with the youth. Often the intervention will focus on issues like anger management, social-emotional competence, and academic performance.

**Environmental focus** (ecological or system-level interventions): Most of these programs modify the social context of the child’s home or school situation (e.g. focuses on parents or teachers or others besides the child).

1. Person centered only (only person)

2. Person centered and one environmental focus (mix)
   [Specify type of environmental change]
   __ School-based
   __ Parent Training
   __ Other: ____________________________

3. Person centered and more than one environmental focus (mix)
   [Specify type of environmental change]
   __ School-based
   __ Parent Training
   __ Other: ____________________________

4. One environmental focus (only environmental)
   [Specify type of environmental change]
   __ School-based
   __ Parent Training
   __ Other: ____________________________

5. More than one environmental focus (only environmental)
   [Specify type of environmental change]
   __ School-based
   __ Parent Training
   __ Other: ____________________________

[conrec] _______ What did the control group receive?
[The difference between ‘received nothing’ and ‘treatment as usual’ hinges on whether or not the two groups have an institutional framework or experience in common, e.g., probation supervision, institutionalization, school, etc.]

1 = Received nothing (no evidence of any treatment or attention)

2 = Wait listed, delayed treatment

3 = Minimal contact, instructions, intake interview, but not wait listed

4 = “Treatment as usual” (TAU)
5 = Attention placebo (control receives discussion, attention, or dilute version of treatment)
6 = Treatment element placebo (Received target treatment except for defined element presumed to be the crucial ingredient)
7 = Weak alternate treatment (control is not really a “control,” but another treatment different than “usual” treatment being compared with the focal treatment; must be a very dilute dose or a “straw man” not expected to perform well)
8 = Substantial alternate treatment other than mental health (same as above except the treatment has sufficient intensity or integrity to be expected to perform well)
9 = Unspecified
For “TAU” or “other”, write in: ______________________________

DURATION and FREQUENCY – Treatment duration in weeks (missing = 999)
Approximate or exact duration of treatment in weeks from first treatment event to last treatment event excluding follow-ups designated as such (divide number of days by 7 and round; multiply number of months by 4.3 and round). Estimate if necessary.
[expdurY] _____ Expected program duration for youth participants(# of weeks) 0 if N/Applicable
[expdurP] _____ Expected program duration for parent/other participants (# of weeks) 0 if N/A
[actdurY] _____ Actual mean program duration for youth (# of weeks) 0 if N/A
[actdurY] _____ Actual mean program duration for parent/other participants (# of weeks) 0 if N/A

[EfforEff] Efficacy or Effectiveness study?
1 = Research team administered the treatment - Efficacy
2 = Non-research team member – Effectiveness

Participant Characteristics
If a study splits the data into groups, e.g., comparison vs. an intervention group, and they don’t separate demographic data for each group, see if the authors checked to see if there were any significant demographic differences between the groups. If there are none, then use the percentages for the entire sample for each group. If there are differences, code the demographic variables were there are
differences as “Missing.” For example, if a study said that the groups differed significantly on age but not on SES, ethnicity, or gender, then use the total sample information for SES, ethnicity, and gender, and code age as “missing.”

NOTE: Treatment and control group should always pertain to the child. Even if a parent/teacher/other received the intervention, the study should provide outcome data related to the child. Thus, for the purposes of this dissertation, the treatment and control group should always be children.

Please use the following formula to calculate percentages:

\[(\text{# of subgroup} \div \text{# of the entire group}) \times 100 = \text{percentage (to 10}^{\text{th}}\text{ decimal place)}\]

Example:

In a sample of 20 children, there were 15 female participants in the study.

\[(15/20) \times 100 = 75.0\% \text{ of sample was female} \]

<table>
<thead>
<tr>
<th>PARTICIPANT CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4_00 1</td>
</tr>
<tr>
<td>What was the number of female participants? [If unknown, enter 999; if youth’s information is not provided separately for treatment and control groups, but the article indicates that both were present, tell Katrina and do not code Section 4 for this article]</td>
</tr>
<tr>
<td>S4_00 2</td>
</tr>
<tr>
<td>What was the number of male participants? [If unknown, enter 999]</td>
</tr>
<tr>
<td>S4_00 3</td>
</tr>
<tr>
<td>What was the total number of participants?</td>
</tr>
<tr>
<td>S4_00 4</td>
</tr>
<tr>
<td>What was the average age of participants? [In years, at start of program, rounded to nearest whole #; use median if average is not available or use average grade level where age = grade + 6. Apply same rule to average age of control youth, min and max age of youth, and modal developmental level of youth. If unknown, enter 999]</td>
</tr>
<tr>
<td>S4_00 5</td>
</tr>
<tr>
<td>What was the minimum age of participants? [If unknown, enter 999]</td>
</tr>
<tr>
<td>S4_00 6</td>
</tr>
<tr>
<td>What was the maximum age of participants? [If unknown, enter 999]</td>
</tr>
<tr>
<td>TR_01</td>
</tr>
<tr>
<td>Predominant race of group. Select the code that best describes the 1 = greater than 60% Whites 5 = greater than 60% 2 = greater than 60% Black 6 = mixed, none more 3 = greater than 60% Hispanic 7 = mixed, cannot 4 = greater than 60% other minority 8 = cannot tell</td>
</tr>
<tr>
<td>What was the race/ethnicity of participants? [Approx. %]</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

**What was the developmental stage of participants? [Approx. %]**

- Early childhood (7 years of age or younger)  
- Middle childhood (8-10 years of age)  
- Late childhood (11-12 years of age)  
- Child unspecified  
- Early Adolescent (13-14 years of age)  
- Middle Adolescent (15-17 years of age; include High School)  
- Late Adolescent (18-21 years of age; include College)  
- Early Adulthood (22-29 years of age)  
- Adolescent unspecified  
- Unspecified  
- Parent/Caregiver

**What was the SES of participants? [Approx. %] Please write in whether context was described. (e.g., 91% low-income school)**

- Low  
- Middle  
- High  
- Unspecified

**What was the educational level of youth? [Approx. %]**

- Preschool  
- Elementary (K-5 or ages 5-10)  
- Middle school (Grades 6-8 or ages 11-13)  
- High School (Grades 9-12 or ages 14-17)  
- Mixed elementary, middle, and high school  
- Post secondary education  
- Technical/Professional school  
- Mixed high school and greater educational level  
- Unspecified

**Risk factors** [Risk factors are those characteristics, variables, or...]

(continues on the next page)
hazards that, if present for a given youth, make it more likely that one or more areas of the youth’s development or adaptation will be negatively affected. When coding any risk factor, endorse only if there is evidence suggesting that it was present in at least 50% of youth.

Did youth exhibit or have in their backgrounds any individual risk factor(s)? Individual risk factors are biological, behavioral, cognitive, or psychosocial characteristic of the youth.

\[ \begin{align*} 
0 &= \text{No} \; \text{(Skip to S4_096)} \\
1 &= \text{Yes} \\
2 &= \text{Unspecified} \; \text{(Skip to S4_096)}
\end{align*} \]

If YES to INDIVIDUAL for youth, specify which factors [Select all that apply. If limited to a variable that can be coded somewhere else, do not code here]

- S4_059 _ Bullying others
- S4_060 _ Fighting and other aggressive behavior
- S4_061 _ Behavior problems at school (other than bullying or fighting)
- S4_062 _ Behavior problems (unspecified)
- S4_063 _ Low academic achievement
- S4_064 _ Truancy/school absenteeism
- S4_065 _ School drop out
- S4_066 _ Learning disorder/disability
- S4_067 _ Intellectual and/or development disabilities
- S4_068 _ Physical disability
- S4_069 _ Poor physical health
- S4_070 _ Mental disorder/mental health problem (internalizing)
- S4_071 _ Depressive symptoms/disorder
- S4_072 _ Anxiety symptoms/disorder
- S4_073 _ Somatic complaints/Somatization disorder
- S4_074 _ Suicidal ideation/attempt(s) (may be related to a variety of disorders)
- S4_075 _ Mental disorder/mental health problem (externalizing)
- S4_076 _ Oppositional defiant disorder
- S4_077 _ Conduct disorder
- S4_078 _ Attention Deficit Hyperactivity Disorder (ADHD)
- S4_079 _ Bipolar Disorder
- S4_080 _ Schizophrenia
- S4_081 _ Psychosis (may be related to a variety of disorders)
- S4_081 _ Mental disorder/mental health problem (unspecified)
- S4_083 _ Early onset of delinquency
- S4_084 _ Favorable attitudes toward delinquent behavior
- S4_085 _ Delinquent behavior (crimes against people)
- S4_086 _ Delinquent behavior (crimes against property)
- S4_087 _ Delinquent behavior (unspecified)
Did youth exhibit or have in their backgrounds any historical risk factor(s)? [Historical risk factors are distal events or experiences that occurred more than a year ago in the youth’s past]

0 = No (Skip to 2 = Unspecified (Skip to S4_111)
1 = Yes

If YES to HISTORICAL factors for youth, specify which factors [Select all that apply]

- S4_104 Family mobility
- S4_105 Non-normative school changes
- S4_106 International immigration
- S4_107 Foster care
- S4_108 Incarceration
- S4_109 Child maltreatment/abuse/neglect
- S4_110 Other(s)

Measures

What educational outcomes were reported (please cite the measure reference)? Check all that apply

[BehVars] (E.g., attendance, misconduct, on-task behavior) 1 = Yes; 2=No

[BehVarsSpec] Specify names of measures __________________________________________
(STRING VARIABLE)

[AchvmtVars] (E.g., GPA, Test scores, Grade promotion/completion) 1 = Yes; 2=No

[AchvmtVarsSpec] Specify names of measures __________________________________________
(STRING VARIABLE)

[AttDsVars] (E.g., academic self-efficacy, school engagement, academic motivation, goal orientation, etc.) 1 = Yes; 2=No

[AttDsVarsSpec] Specify names of measures __________________________________________
(STRING VARIABLE)
[PriMHvar] What is the primary mental health variable. That is, what is the target mental health variable (e.g., anxiety reduction intervention, the primary variable would be one that measures reduction in anxiety)

[ PriMHvarSpec] Specify name of measures _____________________________
(STRING VARIABLE)

Target Mental health outcome[please check]
Psychological

<table>
<thead>
<tr>
<th>1. Depressive symptoms</th>
<th>2. Anxiety symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Mental disorder symptoms mixed</td>
<td>4. Suicide risk</td>
</tr>
<tr>
<td>5. Psychological/emotional distress</td>
<td>6. Psychological/emotional well-being</td>
</tr>
<tr>
<td>7. Coping competence</td>
<td>8. Other, specify:</td>
</tr>
</tbody>
</table>

Antisocial behavior

<table>
<thead>
<tr>
<th>Association with deviant peers</th>
<th>Substance use attitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive/violent behavior</td>
<td>Substance use</td>
</tr>
<tr>
<td>Arrests</td>
<td>Alcohol use</td>
</tr>
<tr>
<td>Delinquency</td>
<td>Other, specify:</td>
</tr>
</tbody>
</table>

Target Academic Outcome[please check]
Academic Outcome(s)

<table>
<thead>
<tr>
<th>1. Classroom behavior</th>
<th>12. School expulsion</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Absences</td>
<td>13. School connectedness</td>
</tr>
<tr>
<td>4. Tardies</td>
<td>15. Academic aspirations</td>
</tr>
<tr>
<td>6. School suspensions</td>
<td>17. Academic competence</td>
</tr>
<tr>
<td>7. Attitudes toward school</td>
<td>18. Achievement – math</td>
</tr>
<tr>
<td>8. School-drop out</td>
<td>19. Achievement Reading</td>
</tr>
<tr>
<td>9. GPA or grades</td>
<td>20. Grade promotion</td>
</tr>
<tr>
<td>10. Grade retention</td>
<td>21. Achievement motivation</td>
</tr>
<tr>
<td>11. General intelligence/Cognitive skills abilities</td>
<td>22. Attitudes towards school/achievement</td>
</tr>
</tbody>
</table>

| 23. Other, specify: |
Effect Size Data

For each effect size, code all of the following items. Note that a study may have subgroups and thus require an Effect Size Form be completed for each individual subgroup.

1. Sample size of intervention and control/comparison group for each time point.
   [Write in appropriate number.]

2. Mean on Educational Outcome measures for each time point
   [Write in appropriate number to the 10th decimal place; Code as “Missing” if data is not provided in the article.]

3. Standard Deviation on Educational Outcome measures for each time point
   [Write in appropriate number to the 10th decimal place; Code as “Missing” if data is not provided in the article.] For each measure please specify whether the measure was subjective (e.g., parent or self-report), objective (e.g., school records), or other.