1-1-2013

Motivational Interviewing and School-Based Mentoring to Improve Middle School Students' Academic Performance

John Terry
University of South Carolina - Columbia

Motivational Interviewing and School-based Mentoring to Improve Middle School Students' Academic Performance

by

John D. Terry

Bachelor of Science
The University of South Carolina, 2008

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Arts in

Clinical-Community Psychology

College of Arts and Sciences

University of South Carolina

2013

Accepted by:

Bradley Smith, Director of Thesis

Mark Weist, Reader

Kim Hills, Reader

Lacy Ford, Vice Provost and Dean of Graduate Studies

ABSTRACT

The development of effective and feasible interventions that are deliverable within schools are badly needed in order to address high levels of unmet academic and social/emotional need in children and adolescents. In order to address these needs, two interventions were developed, delivered, and tested in this study. One was an eight-session School-based Mentoring (SBM) program based on evidence-based academic enabling activities. The other was a one-session report card coaching program based on Motivational Interviewing (MI). Previous studies of these SBM and MI interventions found mostly small or statistically non-significant effects on academic performance. The current study tests the hypothesis that effect sizes may be increased by providing SBM and MI simultaneously, producing an additive or synergistic effect. To address this possibility, a study of the separate and joint effects of the SBM and MI interventions was conducted such that 195 middle school students recruited over two school years were randomly assigned to one of four conditions: SBM only, MI only, SBM plus MI, and a waitlist control group. To implement the SBM intervention, 95 undergraduate students from a southeastern university provided up to seven 45 minute long mentoring sessions. To implement the MI intervention, seven graduate students and three research assistants trained to be "report card coaches" provided one 45 minute long MI session. Specific hypotheses were that MI plus SBM would be superior to waitlist control, and that MI plus SMB combination would be enhanced compared to MI or SBM alone. Results from this study indicate a significant effect for math grades for the MI plus SBM group $d = .28$ but null results for other grades and self-report measures of self-efficacy, life-satisfaction, and school engagement when examining both years of this study combined. However, when examining years separately, in year one there is a slightly higher effect, yet not significant difference, for math grades SBM+MI $d = .38$, SBM $d = .36$, and

MI only *d* = .34, each of these differences were statistically significant from the waitlist control and replicate results from the two previous evaluations of these interventions.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Each day in the United States approximately 21% of the population, nearly 61,530,000 students, can be found within the public school system (www.census.gov/hhes/school/). Of this 21% of the U.S population, approximately 22%, nearly 13,536,600 students, will experience both symptoms and impairment of a psychological disorder (Merikangas et. al, 2010). Psychological disorders most prevalent within children and adolescents are anxiety disorders, disruptive behavior disorders, mood disorders, and substance abuse disorders (Merikangas et. al, 2010). Federal education legislation has categorized students that demonstrate severe emotional and behavioral problems as students with Emotional Disabilities (ED) and these students are eligible for Individualized Educational Plans (IEP) (http://idea.ed.gov). Of students who qualify for an IEP with the disability of having an "emotional disturbance," 52.3 percent of students 14 years old or older dropped out of school (U.S. Department of Education, 2009). For students in general, IES' National Center for Educational Statistics (NCES) indicates a total dropout rate of approximately 7.4% in 2010 (http://nces.ed.gov/pubs2012/2012045.pdf). Some evidence indicates that these statistics may by increasing; Child Trends data indicates that between 2001 and 2007 the number of IEPs for students in kindergarten to third grade increased from 6 to 8 percent of students, meaning 1.2 million more children with IEPs (www.childtrendsdatabank.org). In view of this, a large and seemingly growing number of children and adolescents will experience difficulties achieving academic attainments due to symptoms of behavioral or emotional difficulties.

In addition to students that experience a mental disorder, many students who do not meet

the criteria for an IEP will also experience significant challenges during their scholastic career (Mueller, Phelps, Bowers, Agans, Brown-Urban, Lerner, 2011). As is the case with the majority of students in this study, subclinical behavior problems often times fall within the realm of intervention. For example, the DSM-IV-TR contains V-Codes for common problems that individuals may seek treatment for, but do not meet the criteria of a mental disorder. Scholastic Problems are listed in the DSM-IV-TR as V-code V62.3, which states that youth with "Educational circumstances; academic problem, dissatisfaction with school environment, educational handicap" are common problems that children and adolescents may seek psychological treatment (DSM-IV-TR, American Psychiatric Association, 2000). Given the prevalence of both clinically significant (1 in four) and subclinical problems, there is a substantial need to reach students with effective interventions. However, studies suggest that only one in ten students in need of mental health services receive any services (U.S. Public Health Service, 2000). Moreover, of those who do receive services, most receive inadequate services. According to a frequently cited study of community-based therapy attendance, between 40 and 60 percent of children happen to discontinue prematurely from treatment (Baekeland & Lundwall, 1975; Miller, Southam-Gerow, & Allin, 2008). This may be much shorter than specified by most evidence-based interventions; for example, in some anxiety programs, the mental disorder most prevalent in youth, require 16 sessions to be completed (Podell, Mychailyszyn, Edmunds, Puleo, & Kendall, 2010).

A logical approach to increase the amount of services that students receive would be to offer additional services within schools. Unfortunately, one recent study indicates that between 45.3% of students in need of services receive them (Green et. al., 2013). Thus, a large and potentially growing mental health need exists for youth that school-based interventions could potentially ameliorate. In order to address this need, the Expanded School-mental health (ESMH) movement is underway and aimed at integrating education with comprehensive mental health

services (Waxman, Weist, & Benson, 1999). ESMH desires to promote school community partnerships to provide a broad array of prevention and promotion services to students in both regular and special education (Weist et. al., 2003). Weist (2003) highlights two main themes when encouraging school psychology to increase involvement in ESMH; increased focused on the public health approach and shifting focus toward using evidence-based prevention programs at the population level (Weist, 2003).

This paper evaluates two school-based interventions that are intended to be feasible to implement within a school wide multilevel framework, an intervention structure designed to accommodate student needs in a comprehensive manner along a continuum of intervention strength. First, this paper reviews the rationale and purpose of school wide multilevel systems of support based on public health prevention frameworks. Secondly, we examine how these frameworks serve to provide a context for the selection, implementations, and evaluation of evidence-based practices. While the interventions evaluated in this study are not multilevel interventions themselves they are interventions that could be conceptually be implemented within this framework. Third, we discuss alternative intervention delivery methods, which can serve to ameliorate capacity restraints and expand services via community partnerships and were used in this study of these interventions. Fourth, we review previous studies on the School-based Mentoring and Motivational Interviewing interventions and discuss the importance of replication studies in the process of interventions meeting the criteria for being evidence-based. Lastly, we describe the methods used to evaluate if these interventions do produce similar results and discuss the potential main effects and synergistic interactions between interventions.

1.1 Multilevel Systems of Support

As a result of the significant mental health problems students face, teachers and educational personnel often experience difficulty addressing challenging behavioral and

emotional needs of students. School Psychologists along with other types of interventionists strive for effective and practical interventions that can be delivered within the school walls to address the needs of students. To attend to these issues, School Wide Interventions (SWI) based on the public health prevention model have been applied to schools, most notably in the form of Positive Behavior Interventions and Support (PBIS) (Smith, Molina, Massetti, Waschbush, & Pelham, 2007). This multilevel framework was originally delineated into three main categories: Primary, Secondary, and Tertiary prevention (Walker, & Shinn, 2002; Kloos, Hill, Thomas, & Wandersman, 2012). In this framework Primary Prevention is at the population level and is intended to reduce the rates of new cases; in public health this would be the avoidance of new incidences occurring, whereas prevalence is the current number of cases that are already present in a population. Primary prevention is usually given to everyone in the specified population whether or not they show preliminary symptoms of disorder or not. Secondary Prevention is given to students that are already beginning to show signs of a disorder or are at a particular risk for a disorder. Tertiary Prevention interventions are given to groups that already experience the disorder with the aim of limiting more harm and to prevent further deterioration and exacerbation of symptoms. In 1994, the Institute of Medicine (IOM) proposed another prevention framework that promotes and advances the idea of Universal, Selective, and Indicated Measures for prevention. Universal Prevention is similar to primary prevention in that it is usually given to the entire population. Selective Prevention Measures may be given to people that have increased probabilities for developing a disorder in the future. Indicated Prevention Measures are intended for individuals at higher risk for developing a disorder and may be showing some symptoms but may not be reaching the level for diagnosis. Considerable overlap between levels exists and there is often debate over the discrimination between prevention and treatment, but this prevention model has gained momentum as a school-wide intervention approach.

In schools, prevention efforts are utilized within the context of multi-tiered levels of support aimed at increasing engagement and the frequency of positive behavior in order to decrease the likelihood of inappropriate student behavior. Both the public health and IOM prevention terminology are being applied to school-wide efforts at increasing participation and engagement in programs called Positive Behavior Supports (PBS) or Positive Behavior Interventions and Supports (PBIS). PBIS programs are being implemented in schools to provide programs and services at Tier 1 (prevention and climate enhancement for all students), Tier 2 (prevention and early intervention for students showing early signs of or less severe problems), and Tier 3 (intervention and case management for students showing more significant challenges). PBIS programs are prevention-focused, systematic, and comprehensive in emphasizing depth and quality when addressing student needs in the school building. School-wide PBIS programs are in the process of gaining empirical support with two randomized studies as well as studies examining effectiveness (Chitiyo, May, Chitiyo, 2012; (Bradshaw, Mitchell, & Leaf, 2010; Horner, Sugai, Smolkowski, Todd, Nakasato, & Esperanza, 2009; (Barrett, Bradshaw, & Lewis-Palmer, 2008; Blonigen, B., Harbaugh, W., Singell, L., Horner, R.H., Irvin, L., & Smolkowski, K. 2008).

1.2 PBIS as Context for Evidence-based Practices

Campbell (1986) discusses internal validity of treatment packages in terms of local, molar, and causal validity. Molar validity indicates that treatment packages like PBIS are complex and may contain many potential causes for change. Within a package of PBIS interventions, one cannot be sure which elements of PBIS are working, unless they are tested individually. A challenge is finding interventions to use in the context of PBIS that are effective and ready for dissemination (Chitiyo, May, Chitiyo, 2012). PBIS programs typically include a set of operationally defined core components to guide practice; however, one myth concerning PBIS is that it is an intervention or a package of treatments (Sugai & Horner, 2010). Instead, PBIS is

most accurately described "as a framework or approach that provides the means of selecting, organizing and implementing these evidence-based practices by giving equal attention to (a) clearly defined and meaningful student outcomes, (b) data-driven decision making and problem solving processes, and (c) systems that prepare and support implementers to use these practices with high fidelity and durability" (Sugai & Horner, 2010, p. 4). For example, in PBIS the needs and resources of individual students, families, and classrooms are assessed, and interventions are chosen to target identified needs while simultaneously matching students, families, or classroom strengths. When choosing programs to implement within a PBIS framework, two directories of evidence-based practices that rate their effectiveness and readiness for dissemination exist to facilitate aid in deciding which program to implement. These directories are the U.S. Department of Education Institute of Educational Science's What Works Clearinghouse (WWC) (http://ies.ed.gov/ncee/wwc/) and the Substance Abuse and Mental Health Services Administration (SAMHSA) National Registry of Evidence-based Program and Practices (NREPP) (nrepp.samhsa.gov).

1.3 Alternative Intervention Delivery Methods

Considering the capacity limitations of school mental health workers to reach all of the students in need of mental health supports there is a strong need to provide services in innovative ways (Kazdin, 2011; Prinz & Sanders, 2007). One approach to addressing this issue is task diffusion in which persons with less training provide services within their zone of competency that are usually provided by more highly trained professionals for example nurse practitioners providing some of the services of physicians (Kazdin, 2011; APA, 2008). In schools, task diffusion may include having teachers or paraprofessionals provide services that would usually be provided by specially trained school mental health staff, such as psychologists, social workers, or school counselors (Prinz & Sander, 2007).

Another innovative way to meet the needs of students falls into the category of disruptive innovations. Disruptive innovation “provides a simpler and less expensive alternative that meets the essential needs for the majority of consumers and is more accessible, scalable, replicable, and sustainable” (Rotheram-Borus, Swendeman, & Chorpita, 2012, p 463). For example, providing vaccines in grocery stores is a disruptive innovation that increases the reach of vaccinations to the public and expands the vaccination model from formal medical settings. School staff and community volunteers providing intervention who are not part of the formally recognized school mental health professionals can expand the reach of mental health services in a school beyond formal providers. Task diffusion (i.e., using paraprofessional school based mentors) and disruptive innovations fit within the intervention ecology of schools and provide additional capacity to intervene with youth by providing supplemental personnel to provide interventions within their zone of competency. School-based mentoring provided by college students that meets unmet mental health needs is an example of task diffusion.

1.4 School-based mentoring

Mentoring is a wide spread practice for intervening with youth and is common practice in schools (Randolph & Johnson, 2008). Mentoring is defined

as a "relationship between an older, more experienced adult and an unrelated, younger

protégé—a relationship in which the adult provides ongoing guidance, instruction, and encouragement aimed at developing the competence and character of the protégé"(Randolph & Johnson, 2008, p. 177). There are key distinctions made between community-based mentoring (CBM) and school-based mentoring (SBM) (Grossman, Chan, Schwartz, & Rhodes, 2012). For example, SBM takes place within the school year, occurs within the school facility, and usually has an academic focus. In one large study of mentoring, Grossman et al. followed over 1,139 youth in a mentoring program finding that SBM relationships tended to be much shorter in

duration; most SBM relationships last 6 months on average (Grossman, Chan, Schwartz, & Rhodes, 2012).

SBM is the most popular, most funded, fastest growing, and most studied form of youth mentoring in the United States (Herrera, Grossman, Kauh, & McMaken, 2011). Millions of youth receive mentoring each year within the context of the public school system, and some states have formally incorporated school-based mentoring into Response to Intervention (RTI) and Positive Behavioral Intervention and Support (PBIS) programs (e.g. Dunlap, Goodman, & Paris, 2010). However, data on the effectiveness of SBM is lacking. We found the only form of SBM listed on registries of evidence-based programs is the Check and Connect SBM program, which is described as having limited effectiveness (http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=78). Thus, SBM is an example of a school-based program in need of further study. The overall evidence for mentoring appears to be mixed; only one school-based mentoring program is designated as effective (i.e., Check and Connect). Thus, it seems appropriate that the What Works Clearinghouse site endorses school-based mentoring "with reservations." Clearly, further research and development is needed for SBM.

Over the past three years, two small-scale randomized studies explored the potential of SBM in a local school. The first study examined relationship-based SBM and found disappointing results (McQuillin, Smith, & Strait, 2011). The initially tested version of SBM relied on traditional mentoring practices meaning and interventions based primarily on building relationship between youth and mentors. To address the shortcomings of the relationship-based SBM protocol, the investigators made three major changes to the intervention. First, they added evidence-based academic enabling procedures (i.e., goal setting, organizations skills). Second, they developed much more rigorous training procedures with the incorporation of enhanced supervision. Third, they developed a strong implementation support system with the mentors

checking in with a supervisor before and after each mentoring session. This study found some positive impacts on math grades (*d* = .37) and life-satisfaction (*d* = .39).

1.5 Motivational Interviewing

Motivational Interviewing (MI) is an evidence-based practice originally developed for substance abuse problems that has been generalized to a variety of different needs and settings (nrepp.samhsa.gov; Miller & Rollnick, 2012). MI is a "collaborative, goal-oriented style of communication with particular attention to the language of change. It is designed to strengthen personal motivation for and commitment to a specific goal by eliciting and exploring the person's own reason for change within an atmosphere of acceptance and compassion" (Miller & Rollnick, 2012, p. 29). The developers of MI describe the spirit of MI by featuring core values of evocation, collaboration, and autonomy; the newest edition of *Motivational Interviewing* maintains this description but adds acceptance and compassion. Evocation refers to eliciting motivation to change from the client as opposed to educating them about the need to change. Collaboration refers to the therapist allowing and supporting the client to lead the dialogue and direction of the interaction. Autonomy is the therapist's acknowledgement and support of the client's freedom to choose his or her goals and behaviors. Four main principles of MI are express empathy (i.e., demonstrate accurate understanding of student's thoughts and feelings concerning school), develop discrepancy (i.e., compare student's values concerning school with their current behavior), roll with resistance (i.e., manage resistance in a manner that reduces ambivalence for change), and support self-efficacy (i.e., self motivational statements about change). A primary function of the counselor during MI is to evoke change talk. Four types of change talk have been identified: disadvantages of the status quo, advantages of change, optimism of change, and intention to change.

MI is an approach that is increasing in popularity and in scope of practice with substantial revisions to the first (1992), second (2002), and third editions of the definitive book, *Motivational Interviewing* (Miller & Rollnick, 2012). Given its' efficiency and effectiveness with adults and adolescent substance use, (Leary-Tevyaw & Monti, 2004; McCambridge & Strange, 2004; Kelly & Lapwworth, 2006) it seems reasonable that an approach like MI could be an helpful intervention in academic settings to address scholastic problems and risky adolescent behavior (Naar-King & Suarez, 2011; Frey, Cloud, Lee, Small, Seeley, Feil, Walker, & Golly, 2011; Strait, et al., 2012). However, controlled experimental research in academic settings has only started to appear in empirical peer reviewed publications (Strait, Smith, McQuillin, Terry, Swan, and Malone, 2012). For example, Strait conducted a study that involved 103 students randomly assigned to either a MI (n=50) or "school as usual" control (n= 53) (Strait et al, 2012). Students in the MI condition participated in a single 45-minute MI session. In comparison to the control group, students who received MI reported significant improvements in their class participation and overall positive academic behavior. The MI group also had significantly higher 4th quarter math grades than students in the control group, but effect sizes were small for math (d = .47) and the variance in other grades was not statistically significant. Positive significant effects were also found for self-efficacy and self-reported positive academic behavior (homework completion and participation in class), yet self-reported measures on academic behavior in this study were not significant. In this study, fidelity data indicated that actual sessions ranged from 28 to 63 minutes with the average session being 44.26 minutes. Report Card Coaches reported that they used MI techniques for 95.26% of the core components of the MI sessions. Thus, these initial findings provide tentative support for using MI to improve math grades, but do not meet the standard of being empirically supported.

1.6 Replication Studies

Replication of interventions is a key consideration in documenting their effectiveness and can be considered an ethical imperative (Flay et. al, 2005 ; McFall, 1991). Yet, popularity and dissemination of interventions often precedes their careful replication. McFall (1991) highlights the ethical standards guiding the delivery of psychological interventions to the public: First, the exact nature of the service must be described clearly. Second, the claimed benefit of the service must be stated explicitly. Third, these claimed benefits must be validated scientifically. Fourth, possible negative side effects that might outweigh any benefits must be ruled out empirically (McFall, 1991).

The Society for Prevention Research (SPR) commissioned a task force to develop guidelines for identifying programs that are efficacious, effective, and ready for dissemination (Flay, Biglan, Boruch, Castro, Gottfredson, Kellam, Moscicki, Schinke, Valentine, & Ji, 2005). According to their guidelines, an intervention can be disseminated when it requirements for both the criteria for efficacy and the additional criteria for effectiveness. The SPR task force identified a list of 47 criteria that need to be meet before a program is ready to be disseminated. Five main requirements are identified for an intervention to demonstrate efficacy. Efficacy can be said to be achieved after it has "been tested in at least two rigorous trials that (1) involved defined samples from defined populations, (2) used psychometrically sound measures and data collection procedures; (3) analyzed their data with rigorous statistical approaches; (4) showed consistent positive effects (without serious iatrogenic effects); and (5) reported at least one significant long-term follow-up" (Flay et. al, 2005, p 151). The task force defines a treatment as effective treatment if "not only meet all standards for efficacious interventions, but also will have (1) manuals, appropriate training, and technical support available to allow third parties to adopt and implement the intervention; (2) been evaluated under real-world conditions in studies that included sound measurement of the level of implementation and engagement of the target audience (in both the intervention and control conditions); (3) indicated the practical importance

of intervention outcome effects; and (4) clearly demonstrated to whom intervention findings can be generalized" (Flay et. al, 2005, p 151). However, many disseminated programs lack replication; for example, less than 13% of 895 violence prevention programs had a replication study (Aos, Cook, Elliott, Gottfredson, Hawkins, Lipsey, & Tolan, 2011). Given that this situation is somewhat representative of other areas of intervention, there is a substantial need for school-based replication of initial efficacy studies to establish a dependable set of interventions that are appropriate to use in schools. At this point, neither the SBM intervention nor the MI interventions have met all criteria for efficacy, however replication of originally obtained results serve to meet the replication criteria for efficacy.

1.7 Summary and Study Aims

One important aim of this study is to replicate previous research findings on the SBM and MI intervention developed for middle school students by our research group (i.e., finding the separate effects of SBM and MI). Another important aim of this study is to test for a potential interaction of these two interventions. This is an important consideration that is especially pertinent to multi-modal, multi-level interventions that may use numerous separate interventions. Interventions consist of multiple components that make up a treatment package; this molar approach to conceptualizing interventions raises the possibility of individually validated treatments having joint effects that can interact to strengthen the interventions beyond their separate, additive effects (Campbell, 1986). For instance, SBM and MI could interact to create a powerful synergy that leads to stronger than expected effects of providing either intervention alone or as an additive combination. In this study, SBM was chosen because it is a very popular form of task diffusion and has some preliminary support for efficacy—but with room for improvement (McQuillin et al, under review). MI was chosen because it is effective, efficient, and has preliminary support for effectiveness in schools (Frey et. al., 2011) —and may have the ability to strengthen other interventions. For example, pilot research demonstrates increase effects

when MI is delivered before cognitive behavioral interventions (Kertes, Westra, Angus, & Marcus, 2011; Cornelius, Douaihy, Bukstein, Daley, Wood, Kelly, & Salloum, 2011). The purpose of this study is to test the separate and combined effects of SBM and MI on school behavior and performance in a sample of middle school students.

To the best of our knowledge, this study is the first systematic test of providing SBM and MI simultaneously. Consequently, this design provides a novel test whether there is an incremental benefit of providing the SBM and MI together. Another unique test in this study is to see if there is a joint effect of providing SBM and MI concurrently, but without coordination. Meaning in the SBM and MI group, mentors and MI personnel did not know that students in that group received both interventions and they did not coordinate during their intervention with students. A final consideration in this study is to systematically assess the costs of each intervention, including staff costs, training time and costs, delivery time, and the cost of providing the interventions with fidelity. Comparing these costs and the nature of personnel required to deliver the services is an important feasibility consideration for further use of these interventions.

CHAPTER 2

METHODS

2.1 Research Design

The current study builds on two prior randomized studies conducted through the Department of Psychology at the University of South Carolina (USC) at a local middle school. One of the studies was on School-Based Mentoring (SBM) as compared to control (McQuillin, Smith, & Strait, 2011) and the other was on Motivational Interviewing (MI) as compared to control (Strait, Smith, McQuillin, Terry, Swan, & Malone, 2012). Both of these studies were conducted independently in separate academic years, but with the intention of replicating the MI and SBM interventions in a later year. Accordingly, the manualized procedures for each of these interventions were obtained from the authors and carefully followed (See Appendix A and B). In the current study, a four group randomized repeated measures design was conducted across two years. Students were recruited in the first quarter of the 2011 and 2012 school year and randomly assigned to either: 1) Mentoring plus MI, 2) Mentoring only, 3) MI only, or 4) a waitlist control group. Of note, random assignment was conducted utilizing an online random number generator designed specifically for creating random numbers for scientific experimentation (Urbaniak, & Plous, 2011).

2.2 Recruitment of Middle School Student Participants

Prior to recruitment of study participants, approval was obtained from the USC Institutional Review Board, Richland School District One's Office of Research, Assessment, and Evaluation, and the principle of the middle school. Recruitment of students took place across two years, with similar procedures for both years. Each year, the middle school student participants

were recruited during the first nine weeks of the school year. Approximately 10 research assistants and specialists visited each classroom in the school during first period to make an announcement about the mentoring and report card coaching program. These research staff were given a brief script to read before visiting classes and encouraged to use the script verbatim. After making the announcement describing the study, any student that expressed interest in participating was given a consent form to take home to obtain a parent signature. Students were asked to return the consent forms to their first period teacher or to take the forms directly to the attendance office and give them to the school's contact person for the study. The only difference in recruiting was that during the second year of the study, additional copies of consent forms were placed into teachers' mailboxes and an additional announcement was made during the school wide morning announcements and parent voicemail system. Timeline for recruitment was similar across years one and two, occurring from the fifth week of the first academic quarter and lasting for three weeks.

After students returned their consent forms, research staff removed them from one elective class in order for assent forms to be explained and reviewed with the participants. After assent forms were explained and signed, pretest surveys were administered. During this time students were also asked to volunteer to miss the elective class for mentoring or report card coaching once a week for a semester. Students were not allowed out of core academic classes or electives receiving high school credit. All assessment and intervention was performed during the elective periods of the school day.

After the initial recruitment process for year one, 98 students were included in the total sample. However, two students did not participate in the study because they received mentors from another community organization that was not following the SBM protocol. The decision to not allow these students to participate was made by the school contact person for the study. The first year sample size of 96 middle school students participated in the study. During Year One of

the study, after random assignment to intervention groups, there were 21 middle school students receiving Mentoring, 25 receiving MI, 26 receiving combined intervention, and 24 serving in the waitlist control condition. Demographic information for this sample consisted of 43.75% 6th graders, 29.17% 7th graders, and 27.08% 8th graders along with 81.25% African American, 14.53% Caucasian, and 0.04% Hispanic. Gender consisted of 44.79% Male and 55.20% Female with 62.50% receiving Free Lunch 07.29% receiving Reduced Lunch, and 26.04% receiving Unsubsidized Lunch (See Table 1 for demographic information for students in year one). During the Year Two of the study 99 total students were recruited, after random assignment to intervention groups, there were 27 middle school students receiving Mentoring, 22 receiving MI, 26 receiving combined intervention, and 24 serving in the waitlist control condition. Demographic information for this sample consisted of 57.29% 6th graders, 34.37% 7th graders, and 11.46% 8th graders along with 77.08% African American, 22.92% Caucasian, and 0.00% Hispanic. Gender consisted of 43.75% Male and 59.37% Female with 57.29% receiving Free Lunch 10.42% receiving Reduced Lunch, and 26.04% receiving Unsubsidized Lunch (See Table 2 for demographic information of students in year two). For both years combined, after random assignment to intervention groups, the total sample across two years consisted of 48 middle school students receiving Mentoring, 47 receiving MI, 52 receiving combined intervention, and 48 serving in the waitlist control condition. In this sample there were 97 sixth graders, 56 seventh graders, and 52 eight-grade students. Demographic information for this sample consisted of 49.74% 6th graders, 31.28% 7th graders, and 18.97% 8th graders along with 77.94% African American, 18.46% Caucasian, and 0.02% Hispanic. Gender consisted of 43.59% Male and 56.41% Female with 58.97% receiving Free Lunch 08.72% receiving Reduced Lunch, and 25.64% receiving Unsubsidized Lunch (See Table 3 for demographic information of students across years one and two).

2.3 MI Procedures

In the MI intervention, students were assigned a graduate student or research specialist (i.e., an advanced undergraduate or recent graduate) who was trained to serve as a "Report Card Coach" using the manual and procedures developed by Strait et al. (2012). Report card coaching was provided to a total of 25 middle school students during year one and 22 student during year two, both receiving only one session during the second nine weeks. These sessions were intended to last approximately 45 minutes, which is equivalent to the time of a typical middle school class period. As part of a self-report fidelity checklist, Report Card Coaches timed how long each session lasted, the mean MI time for year one was M= 39.30 minutes with SD= 6.04 minutes and ranged from 23 to 50 minutes and for year two was M= 42.60 minutes with SD= 5.23 minutes and ranged from 31 to 47 minutes. Report Card Coaches report that 98.65% of the core components of the intervention were completed for year one and 97.23% for year two. There were 92 total sessions (combining MI group and the MI+SBM group), thus students were pulled out of class for a total of 92 times resulting in approximately 4,232 minutes of total MI intervention across both years. As mentioned previously, Report Card Coaching sessions were conducted at a time that did not interfere with core academic courses. MI procedures were the same for each student regardless of age or grade.

2.4 Reactance and Empathy Measures

As an additional form of fidelity monitoring, students completed self report measures of empathy and reactance directly after the report card session occurred. In the Strait et. al. (2012) study, two versions of the Motivational Interviewing Reactance Scale (MIRS) scale (Strait et. al., 2002) were developed to specifically measure resistant behavior of the student and resistant eliciting behavior of the report card coach. The questions pertaining to the student's resistance were developed using Miller and Rollnick's four process categories of client resistant behavior (Miller & Rollnick, 2002) The client-form (MIRS-C) measured the student's perspective of the MI session and the interviewer-form (MIRS-I) measured the interviewer's perspective of the MI

session. All fifteen questions were answered using a six-point scale ranging from “never” to “very often”. Specific items were: *When talking to the student I argued that he or she should change his or her academic behavior, I lectured the student about changes he or she should make in school, I asked a lot of "yes" or "no" questions, I tried to make the student feel bad for their current academic standing by criticizing and shaming their current behavior, I gave the student a label (e.g., troubled) to explain their behavior, I appeared to be very rushed or acted as if I only had a few minutes to talk, I was bossy or acted as if I knew what was best for me for the student, The student argued that they did not need to change their current behavior in school, The student interrupted me during our discussion (i.e., talked over or cut off), During the discussion with the student expressed that he or she did not feel responsible for their academic achievements and misfortunes (e.g. blaming), The student did not pay attention to my questions or statements, The student did not respond to my questions, The student tried to change the subject of the conversation, During the discussion the student made excuses for their current academic situation, The student expressed that he or she did not need to change or follow their change plan.* Total reactance scores were obtained by summing the ratings for each question meaning, higher scores indicate higher levels of reactance. In the Strait et. al. (2012) study, the MIRS was completed by all students in the treatment group ($\alpha = .75$) and interviewers following each individual MI session ($\alpha = .67$). The interviewers’ reports of reactance was not significantly related to the students’ report of reactance r (48) = -.20, p = .163; indicating small and negative agreement between the interviewer and student. The MIRS has total sum score of 90, with lower responses indicating lower levels of reactance was elicited during the session. In the current study, the MIRS was completed by all students in the treatment group ($\alpha = .79$) M = 26.43 SD = 10.34 and report card coach following each individual MI session ($\alpha = .76$) M= 19.91 SD = 5.66. The interviewers’ reports of reactance were significantly related to the students’ report of reactance r (46) = .42, p = .004; indicating a medium sized (See Cohen for correlation size guidelines) agreement between the interviewer and student (Cohen, 1992).

The second self report measure completed by both students and report card coaches immediately after the MI session was the Consultation and Relational Empathy Measure (CARE). This measure was originally developed to measure patient's perspective of their medical doctor's ability to express empathy (Mercer et al., 2004) and was adapted for the Strait et al study. For this study the ten-question survey was slightly modified to measure the student's perspective of the report card coach's ability to express empathy during the MI session. Specific items were: *How was I at making the student feel at ease, How was I at letting the student tell his or her "story", How was I at listening to the student, How was I at being interested in the student as a whole person, How was I at fully understanding the students concerns or goals or values, How was I at showing care and compassion, How was I at being positive, How was I at explaining things clearly, How was I at helping the student to take control, How was I at making a plan of action with the student.* We also developed an interviewer-form (CARE-I) in order to measure the interviewer's perspective of their own ability to express empathy. All questions were answered using a six-point rating scale that ranged from "poor" to "excellent" and "does not apply". The CARE has total sum score of 60, with higher responses indicating higher levels of empathy expressed during the session. In the Strait et al. study, the CARE was completed by students in the treatment group ($\alpha = .84$) and interviewers ($\alpha = .90$) immediately following each MI session. In this study, the interviewers' report of empathy (based on CARE) was not significantly related to the students' report of empathy, $r\ (48) = .14$, $p = .332$; indicating low agreement between the interviewer and the student. In the current study, the CARE was completed by students in the treatment group ($\alpha = .93$) M= 47.14 SD = 5.81 and interviewers ($\alpha = .76$) M= 42.91 SD = 5.20 immediately following each MI session. In this study, the interviewers' report of empathy (based on CARE) was not significantly related to the students' report of empathy, $r\ (87) = .09$, $p = .418$; indicating low agreement between the interviewer and the student.

2.5 MI Personnel

Training procedures were closely followed from the original Strait et al. study (Strait et. al., 2012). For Year One, three graduate students enrolled in a Clinical-Community Psychology doctoral training programs and three bachelor-level research specialists served as Report Card Coaches. These graduate students had a mean of approximately 3 years of graduate training experience; the research specialists had a mean of approximately 1.5 years experience providing direct services with youth in an evidence-based after-school program, in addition to their exposure and training with basic counseling skills. For year two, four graduate students enrolled in the either a Clinical-Community or School Psychology doctoral training program and two bachelor-level research specialists served as Report Card Coaches. These graduate students had a mean of approximately 3.25 years of graduate training experience; the research specialists had a mean of approximately 2.0 years experience and exposure and training with basic counseling skills.

2.6 MI Training

The principal investigator and his research advisor conducted biweekly trainings for the Report Card Coaches. The PI's research advisor coauthored the MI protocol from the previous study and had participated in previous MI trainings. The PI of this study had delivered approximately twenty-five academic coaching sessions during the Strait et al. study (approximately half of the entire intervention), and had other training and experience in motivational interviewing. Owing to this experience, the PI was able to provide training and supervision in the delivery of the manualized semi-structured interview with minimal assistance of the faculty research mentor.

During the three training sessions, MI skills were taught through handouts, didactic instruction, discussion, practicing skills through role-playing scenarios, and debriefing in order to give feedback about performance. Each Report Card Coach was required to demonstrate mastery

of the protocol and MI skills in order to begin intervention with participants in the study. Each Report Card Coach achieved greater than 95% fidelity during a final accreditation role-play.

2.7 SBM Procedures

In the school-based mentoring program, middle-school students were assigned a mentor from an academic magnet program of a large southeastern university. These mentors met with students once a week during the school day at a time that did not interfere with academic instruction. All of the mentoring activities were on school grounds during school day. During mentoring meetings, the students were allowed to talk about whatever they choose; however, each session was structured such that mentors followed a daily checklist in the mentoring manual and were instructed to complete all the core components for each session (See Appendix A). The core components of each session were: completing homework and agenda checks, homework and test preparation assistance, organization and planning checks, and completing a weekly academic goal (See Appendix A). This was considered mentoring, rather than tutoring, because the focus was on the process of academic planning and goal setting as opposed to assisting with specific academic content (as would be the case in tutoring). SBM procedures were the same for each student regardless of age or grade.

2.8 Recruitment of Mentors

Mentors were recruited from a university magnet program for students with high tests scores and interests in service, leadership, and study abroad. During first week of orientation, a presentation was given to approximately 500 first year students in the magnet program for year one of the study. This was repeated during the second year, but the presentation was given to approximately 600 undergraduates. Additionally, an in class presentation was delivered to a freshman orientation class called “U101” that included approximately 20 students in the magnet program across both years. After these initial presentations, follow up informational meetings

were offered to students interested in participating in the mentoring program. During these meetings, the participating students filled out background checks and applications and completed pretest service-learning surveys. During these meetings in year one, 63 mentors were recruited, however only 42 mentors successfully completed all necessary paperwork, trainings, and were able to accommodate the limited middle school mentoring times in year one. In year two, 57 mentors were recruited; however, only 51 mentors completed the training and paperwork requirements. By participating in this program, mentors were given credit for fulfilling a community service of their magnet program. The average age of these mentors was approximately 18 years with 23.91% male and 76.09% female.

2.9 Mentor Training

Over the course of 5 weeks, mentor trainings were offered on various days of the week at a variety of times during the day and at night (See training timeline, Appendix B). The PI and a research assistant for the study conducted the trainings, each of which lasted approximately two hours. Each of these trainings consisted of instruction on the mentoring manual, modeling of mentoring skills, and behavioral rehearsals of common mentoring scenarios. Undergraduate students that filled out paperwork but did not attend the first mentor training session were not allowed to participate in the program. As a final preparation to mentoring, "dry run" visits to the middle school were conducted in order for the mentors to arrange transportation, rehearse arriving at the school on time, signing in at the attendance office, and checking in with the on-site mentoring contact person. Thus, the mentors spent about 5 hours preparing to do the mentoring.

2.10 Mentoring Sessions

The mentoring manual describes detailed procedures for each mentoring session and includes a session checklist for key components of each session (See Appendix A). Mentors were instructed to bring their manual each day, review the session before meeting with their protégé,

review the checklist fifteen minutes before completing the session without directly letting the mentee observe them reviewing the checklist, and make pertinent session notes after each session. When arriving at the mentoring room each day, a research assistant asked the mentors to describe the plan for that mentoring session including specific objectives. At the end of each session mentors were asked to self-report on what was accomplished that day. During Year One mentors completed M = 5.50 SD = 1.78 sessions with a range of 0 to 7 session. During Year Two mentors completed M = 5.12 SD = 1.60 mentoring sessions with a range of 0 to 9 sessions. For both years combined, Mentors completed M = 5.33 SD = 1.70 mentoring sessions with a range of 0 to 9 sessions.

2.11 Motivational Interviewing Plus Mentoring

In order to examine the potential combined effects of MI plus SBM, students were randomly assigned to a condition receiving both interventions. These interventions were crossed where students in this group received both interventions. In order to prevent cross-intervention contamination, Mentors and Report Card Coaches were trained separately, were blind to which students were receiving both interventions, and did not share any information with each other when conducting sessions with students. In order to protect against any type of contamination effect, only the primary investigator and one research assistant were aware of student group membership.

Based on the above, the MI and SBM interventions were separate and distinct interventions. We considered creating a third intervention that was a combination of MI and SBM in which the mentors and MI providers coordinated their efforts. However, an important aim of this study was to provide exact replications of previous protocols, measuring the separate effects, and measuring the incremental or synergistic effects of SBM and MI. If we had coordinated these

two interventions we may have created an entirely new intervention, making comparison with previous studies difficult.

2.12 Measures

Data on students' self-reported academic behavior, affective and behavioral school engagement, self-efficacy, and life satisfaction were collected for this study. The CARE empathy and MIRS reactance measures discussed above were intended to be process measures that were completed by participants and Report Card Coaches at only one time point; therefore they are reported under the procedures section and not included in these analyses. The measures below were chosen in order to guide future studies examining the mechanisms of action of these interventions, however, at this point in the evaluation of these interventions these constructs are theoretically interesting, but exploratory. Pretest measures (self-reported academic behavior, affective and behavioral school engagement, self-efficacy, and life satisfaction) were given six weeks into the first quarter and took approximately four weeks to complete. Posttest measures (repeated administration of the same survey measuring the constructs above) were administered during the 17th week of school (two weeks before the second nine weeks was completed). Students participating in the study were asked to complete pretest and posttest surveys under the supervision of research assistants.

The APA task force on statistical inference provides guidelines for statistical methods and the reporting of measures in psychology journals, these guidelines encourage authors to report as much information as possible concerning the psychometric performance of their measures, especially when investigating novel interventions or when performing research in new areas (Wilkinson, 1999). The following analyses are meant to be descriptive, serve to be diagnostic in nature, and provide as much information as possible concerning how these measures performed.

2.13 Academic Self-efficacy

Academic self-efficacy sub-scale items were created specifically to match target skills taught during these interventions. These items were developed based on recommendations by Pajares and Miller for measuring self-efficacy in students (Pajares & Miller, 1995). In this approach, students rate their certainty in their ability to successfully perform an academic behavior by endorsing a 1 to 10 Likert scale (1 = not at all certain to 10 = very certain). The first five self-efficacy items were developed by Strait (2012) for the original study of MI with middle school students (Strait et al., 2012). These items contained content concerning student's belief in their ability to: *complete homework, ask teachers questions, take notes and participate, listen without getting off task, and earn the AB honor roll*. These items were of specific interest to the MI intervention, since increasing self-efficacy around these behaviors is a focus of the MI protocol. In the Strait et. al. study, the self-efficacy scale demonstrated poor levels of internal consistency ($\alpha = .61$). Strait et. al. (2012) used the control groups' pre-test and post-test scores of self-efficacy to estimate test retest reliability, this procedure is followed here to aid comparison of measures across studies (See Strait et. al., 2012, Terry et. al., In press). Pearson Correlations indicated moderate test retest reliability of overall self-efficacy, r (51) = .58, $p < .001$. This procedure for calculating test retest reliability is not a traditional means of estimating test retest reliability; test retest reliability is usually established during scale development. However, in the spirit of the APA task force's recommendations on reporting psychometric properties and given that these items were developed specifically for the MI intervention, test retest reliability was calculated using the control group pretest and posttest measures. These calculations of test retest should be interpreted with caution, yet may provide some beneficial information concerning scale performance.

In the current study, three additional academic self-efficacy items were added concerning student's ability to: *organize their locker, use an agenda book effectively, and make changes*

needed to reach academic goals. These items were added to the measurement packet since the first two academic skills were targeted for intervention in the SBM protocol and the third question was of specific theoretical interest in the MI protocol. Meaning, during each SBM session mentors would perform locker and agenda book checks to ensure that students were organized and correctly filling out their agenda book. The third item was intended to target students' self-efficacy at making progress on their report card coaching goal.

Because there were different items used across the original Strait MI study, McQuillin SBM study, and the current study items; three scales were created and analyzed separately. The first scale averaged the five items from the original MI study. At pretest assessment across both years in the current study, this scale has an internal consistency reliability of (α = .48). The second scale averaged three items that were targeted for change in the SBM intervention (α =0.56). A third total overall self-efficacy scale was created with all nine self-efficacy items (α = 0.63). These alphas suggest poor internal consistency reliability and that these scales should not be used.

Pearson correlations using data from the control group indicated poor test retest reliability of the nine-item scale overall academic self-efficacy r (41) =. 15, p = .35. Pearson correlations, using the entire sample, between pretreatment overall self-efficacy and mean grades for four core classes was also low in the small range: r (185) = .27, p < .001. See Table 6 for test retest reliability for the overall self-efficacy scale and ecological validity (Pearson correlations) of this scale to mean grades and each core subject area. Pearson correlations were used to examine ecological validity, the extent to which procedures and measures in this study match meaningful real world measures (Schmuckler, 2001). Additionally, test retest reliability was computed for each of these items separately (See Table 4) as well as the ecological validity was computed for each item to each core subject area (See Table 5). Pearson correlations were examined individually to assess which items were most closely associated with the dependent variable of

interest, loosing implying which items may be performing best in terms or internal consistency reliability (again, these statistics are meant to be descriptive) (Cortina, 1993) .

2.14 Positive Academic Behavior

Students were asked to rate their current academic behavior in several key areas on a Likert scale ranging from 10% of the time to 100% of the time. Student rated their academic behavior on the following items: *how often do you participate in class, turn in homework assignments, how often are you called down by a teacher, how many teachers have called you down in the past two weeks, and how much time have you spent participating in youth activities*.

Strait et al. (2012), reported results on overall positive academic behavior by averaging two items, *percentage of the time students turn in homework* and *percentage of the time student participates in class*. Following data analytic procedures from the Strait et al. study, two items (percentage of time participating in class and percentage of homework competed) were averaged together in order to create an overall positive academic behavior scale. In the Strait et. al. study, at pretest assessment the reliability of the positive academic behavior scale was calculated to be poor ($\alpha = 0.64$) (i.e., average score between participation item and completed homework item). In the current study at pretest assessment the internal consistency reliability of the positive academic behavior scale was calculated to be poor ($\alpha = 0.57$). This suggests that the items have little internal consistency and should not be used.

In the Strait et. al. study, Pearson correlations indicated that self-reported pretreatment participation, r (101) = .20 to .34, $p < .05$, and homework, r (101) = .40 to .51, $p <.001$, both significantly related to pretreatment grades; providing some ecological validity of this measure. The control groups' pre-test and post-test scores of participation and homework completion were used to estimate test retest reliability. Pearson Correlations indicated moderate test retest reliability for participation, r (51) = .55, $p < .001$, and homework, r (51) = .46, $p < .001$. In the

current study, Pearson correlations indicated moderate test retest reliability for participation, r (47) = .44, p = 0.04, and homework, r (47) = .43, p = 0.005. Pearson correlations, using the entire sample, between participation and mean core grades indicated small ecological validity, r (185) = .26, p < 0.001. Correlations between homework and grades were found to be moderate, r (193) = .51, p < 0.001. Overall positive academic behavior (i.e., average of homework and participation) has moderate test-test reliability r (41) = .45 p = 0.003 and was also moderately correlated with pretest core grades r (185) = .47, p < 0.001. See Table 7 for test retest reliability for the positive academic behavior scale and ecological validity (Pearson correlations) of this scale to each core subject area. Additionally, test retest reliability was computed for each of these items separately (See Table 4) and ecological validity was computed for each item to each core subject area (See Table 7).

2.15 Life satisfaction

During the McQuillin et al. (2013) SBM study, students were asked to complete the Brief Multidimensional Student's Life Satisfaction Scale (BMSLSS). This seven-item scale asked students to rate the level to which they agree with statement concerning their satisfaction with life (Huebner, 1991; Greenspoon & Saklofske, 1998). Students endorsed responses ranging from Strongly Disagree to Strongly Agree for the following items: *My life is going well, my life is just right, I would like to change many things in my life, I wish I had a different kind of life, I have a good life, I have what I want in life, my life is better than most kids'*. Psychometrics for the McQuillin study were not reported, however during the scale development these 7 items demonstrated acceptable internal consistency reliability (α = .80). In the current study, the life satisfaction scales processed an internal consistency reliability of (α = 0.79) at pretest measurement. Pearson correlations using data from the control group indicated high test retest reliability of life satisfaction r (39) = .71, p < .001. When using the entire sample, Pearson correlations between pretreatment life satisfaction and mean core grades indicate there is

practically no correlation and no ecological validity between life satisfaction and grades in this sample, r (188) = .03, p = 0.70. For each grade separately Math r (179) = .01, p = .94, History r (179) = .06, p = .43, ELA r (179) = -.02, p = 0.81, and Science r (179) = .06, p = .47.

2.16 School Engagement

New to this study, students were asked to complete the Student Engagement and Motivation Questionnaire (SEMQ), which contains subscales for Affective Engagement, Behavioral Engagement, and Perceived Competence. These scales were added to replace a school engagement scale that demonstrated poor psychometric properties during the McQuillin study. The remove measures of connectedness were from the Hemingway Measure of Adolescent Connectedness version 5.5; the 6 item subscale of connectedness to schools (α = .82) and 6 item connectedness to teachers scale (α=.79) (Karcher, 2003). On these scales, students were asked to rate the degree to which they agree to the statements on a one to five Likert scale ranging from Strongly Disagree to Strongly Agree. The affective engagement scale contains items concerned with student ratings on the following items: *I like my school, I am proud to be at this school, I look forward to going to school, I am happy to be at this school, When I'm in class I feel good, When we work on something in class I feel interested, class is fun, I enjoy learning new things in class, When we work on something in class, I get involved.* This eight-item scale possessed an internal consistency reliability of (α = 0. 87). Pearson correlations using data from the control group indicated good test retest reliability of affective engagement r (40) =. 68, $p < 0.001$. Pearson correlations, using the entire sample, between pretreatment affective engagement and core grades were, r (191) = .08, p = 0.27. For each grade separately Math r (191) = .08, p = .28, History r (191) = .09, p = .24, ELA r (191) = .00, p = .98 and Science r (191) = .10, p = .17. The behavioral engagement scale contained the following items: *I try hard to do well in school*; *If I have trouble understanding a problem, I go over it again until I understand it*; and *In class, I work as hard as I can*. (α =0.77). Pearson correlations using data from the control group indicated

moderate test retest reliability of behavioral engagement r (40) =. 55, p < 0.001. Pearson correlations, using the entire sample, between pretreatment affective engagement and core grades indicated weak ecological validity, *r* (183) = .09, *p* = 0.25. For each grade separately, Math *r* (182) = .08, *p* = .26, History *r* (182) = .11, *p* = .14, ELA *r* (182) = .05, *p* = .52 and Science *r* (182) = .04, *p* = .56. The perceived competence scale contained five items: *I am good at schoolwork; I am just as smart as others; I can remember things easily; It is easy for me to make friends;* and *Most kids like me* (α = 0.74). Pearson correlations using data from the control group indicated moderate test retest reliability of perceived competence *r* (40) = .64, *p* < 0.001. Pearson correlations, using the entire sample, between pretreatment affective engagement and core grades indicated weak ecological validity, *r* (183) = .03, *p* = 0.70. For each grade separately Math *r* (191) = .00, *p* = .96, History r (182) = .09, *p* = .23, ELA *r* (182) = -.04, *p* =.60 and Science *r* (182) = .05, *p* = .54.

CHAPTER 3

RESULTS

3.1 Checking of Assumptions

Descriptive statistics and checking of assumptions were conducted for each of the respective statistical tests. Descriptive statistics were used to examine distributional properties and evaluate compliance with the assumptions of the models. Skew and Kurtosis were examined for ELA, Science, Math, History, the self-efficacy subscales, positive academic behavior, life-satisfaction, affective engagement, behavioral engagement, and perceived competence; all absolute values of skew were below 2 and kurtosis statistics were below 3 and were not regarded as severe deviations from normality. Descriptive statistics for pretest and posttest survey variables and grades are found in Table 8. A multivariate analysis of variance (MANOVA) test was conducted on the core academic classes in order to test pretreatment equivalence; no significant group differences existed at pretest.

3.2 **Grades**

Planned comparisons were conducted for SBM+MI vs. waitlist Control, SBM vs. waitlist Control, and MI vs. waitlist Control on each core academic grade. These orthogonal comparisons were planned in advance, therefore no post hoc corrections are need to control for experiment wise error and Type I Error rate inflation (Peugh, 2010). Two-level Hierarchal Linear Models (HLM) were used to analyze the effect of the treatment on post-test grades (i.e., second quarter) after controlling for pre-test grades and other covariates that were significantly correlated to pretreatment grades (See Table 9). HLM analyses were conducted in order to address the non-

independent data structure of the dependent academic grade variable. To control for differences in teacher grading, we added a random effect to control for clustering within classes because interclass correlations (ICC) exceeded recommended levels (ICC = .16 to .27).

For both years of the study, a statistically significant main effect was found for the treatment on post-test math scores for the MI+SBM intervention, B = 2.52 (1.28), t = 1.98, $p < .05$, $d = .28$. This suggests that participants in the treatment condition scored 2.52 grade points higher (scale from 1 to 100) on second quarter math grades in comparison to the waitlist control condition after controlling for pre-test math grades. There were no significant effects for the other groups versus control or significant group differences between treatments. No other significant effects were found for grades, with the caveat that science grades are (currently) untested because the HLM model failed to converge when year 1 and year 2 data were pooled. This is most likely a result of some science classrooms containing only a few students, resulting in the computer programs used to analyze this model (both SAS and R) being unable to produce parameter estimates. All other models converged appropriately.

3.3 Cohort Effects on Grades

During this analysis with two years of data pooled together, results were not statistically significant except math grades for the combined group (i.e., MI+SBM) on math grades relative to controls. However, when year one of these data were analyzed, the HLM analysis conducted on math grades indicate that there were significant group differences between the MI+SBM group, the MI group, and the SBM group versus waitlist control. While these groups are not statistically different from each other, Cohen's d calculations are SBM+MI $d = .38$ compared to SBM $d = .36$ and MI only $d = .34$, each of these differences were statistically significant from the waitlist control. Additionally, science grades for the SBM+MI versus control were found to be significant

with *d* = .35 and SBM only approached statistical significance *d* = .27. In Year One, no significant effects were found for ELA or History grades (See Table 11).

3.4 Academic Behavior and Psychosocial Measures

Multiple regression models examined the effects of the inventions on self-report measures of academic behavior, self-efficacy, and life satisfaction, affective engagement, behavioral engagement perceived competence after controlling for pre-test academic behavior and any predictor variable that was significantly related to the dependent variable of interest (i.e., socio-economic status, age, ethnicity, and honors class status). The effect sizes for the self-report academic behavior and psychosocial measures were small and non-significant.

CHAPTER 4

CONCLUSION

This study was designed to examine the separate and joint effects of the SBM and MI interventions, using data from 195 middle school students randomly assigned to one of four conditions: SBM only, MI only, SBM plus MI, and a waitlist control group. This study builds on two previously developed interventions based on the SBM procedures of McQuillin et al. (2012) and the MI intervention based on the procedures of the Strait et al. (2012). Two preliminary studies of these interventions provide support for the efficacy of MI and SBM as interventions to improve math performance in middle school students. The current study found a significant effect for math in the MI+SBM group and found significant effects for MI+SBM, SBM, and MI in Year One of this study for math. However, given the results seem to be unstable from year to year it is uncertain if there is any benefits of combining MI with SBM. Taken together, these findings along with the Strait et. al. and McQullin et. al. studies provide preliminary support for SBM and MI interventions with middle school students in an effort to improve math grades, however further investigation is needed for the MI+SBM intervention.

This study attempts to build on previous studies examining two novel MI and SBM interventions. It aims to replicates results of two previous. Unfortunately, dissemination of interventions often proceeds developing a strong evidence-based. According to guidelines of program dissemination, an intervention is not ready for dissemination until it meets requirements for both efficacy and effectiveness. Replication is required for both and is one of the most important steps in establishing an empirically supported treatment (Flay et. al., 2005; Valentine et al., 2011). For math grades, during Year One of this study, similar significant effects where found

as McQuillin and Strait, however comparisons across studies remain difficult due to poor measurement of fidelity to procedures. Additionally, this study builds on prior research by comparing three competing interventions, which allows for stronger inferences about respective effectiveness as opposed to only comparing interventions to a no treatment control.

The current study has several major methodological limitations that threaten internal validity and generalization to other settings. First, as was the case with the original Strait et al. (2012) and McQuillin et al. (2012) studies, academic grades as well as several constructs of theoretical interests were measured, but only a handful of these variables were found to be significant. While effects for math were found to be significant in the Strait and McQuillin studies and the first year of this study, this was the only effect found to be significant fairly consistently. It is unclear why these interventions appear to be affecting math but not other academic areas. Owing to limited experimental control and poor measurement of fidelity it is unclear whether the intervention itself differentially effects math grades or variation with the interventions themselves explains this result. A major assumption of experimental techniques in social science is that procedures will be delivered systematically and identically. Even with randomized studies, when interventions are not followed exactly, so called "broken randomized experiments," there is little ability for inference (West, 2010, p 19). This study uses an intent to treat design whereas each student was assigned to a group and then analyzed the same regards if they received less than full intervention. This design is a more conservative test of the interventions' treatment effects, however specific information considering dose-response is lost.

Secondly, middle school students completed self-reports of academic behavior. Several of these measures used in this study demonstrated poor psychometric properties in terms of internal consistency reliability, test retest reliability, and ecological validity (Pearson Correlations between academic grades and measures used in this study). The internal consistency reliability of these measures ranged from $\alpha = 0.48$ to 0.87, indicating that some of these scales were not

measuring the same construct of interest. Future studies should employ measures that psychometrically sound with an emphasis on choosing constructs that aid in the investigation of potential mechanisms of action and are aimed at developing stronger theories of how these interventions may produce change. Additionally, self-reports of behavior are problematic because participants may know what is being measured. Meaning some of the items in this study have high content validity (i.e., how important is it for you to make good grades) resulting in an increased chance of biases responding. Furthermore, in the case of children and adolescents' behavior, self-reports are best used in the context of other information, such as parent ratings, teacher ratings, and objectively observed behavior. Teacher ratings and classroom observations would substantially improve this line of research. Having teachers who are blind to the intervention complete ratings on participation and homework completion would be a major methodological step forward. Some of these data may be readily accessible because most teachers give students grades for participation and homework completion, which are ecologically valid measures of positive academic behavior. However, the precision and validity of teacher ratings should be substantiated by direct observation. Future research studies have the opportunity to increase confidence in previous findings by adding multiple measures of multiple constructs of interest. Additionally, measures with both well-established nomological networks (i.e., construct validity evidence) and ecological valid measures that correspond with measures that are germane in real world applications (e.g., parent report, teacher report, standardized testing data, and objective fidelity measurement) would serve to substantiate current self-report data from students (Cronbach & Meehl, 1955; Schmuckle, 2001).

A final consideration was that this program of research has, so far, relied on psychology graduate students (and highly trained recent graduates), as opposed to school personnel, to provide MI. The supply of university students to provide MI is limited, thus threatening the reach of school-based MI. To address this issue, future studies should address the acceptability and

feasibility of recruiting other providers to implement MI. This may include school personnel (including teachers, school administrators, school counselors, school mental health staff), or paraprofessional volunteers to provide MI. Future studies should also address the practical issues (i.e., acceptability, feasibility, and sustainability) of combining MI with SBM. This may include having SBM and MI provided by the same person, or having a coordinator work with separate mentors and MI providers to coordinate their efforts.

REFERENCES

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC, *American Psychiatric Association*, 2000.

American Psychological Association Task Force on Evidence-Based Practice for Children and Adolescents. (2008). *Disseminating evidence-based practice for children and adolescents: A systems approach to enhancing care.*Washington, DC: American Psychological Association.

Aos, S., Cook, T., Elliott, D., Gottfredson, D., Hawkins, D., Lipsey, M., Tolan, P. (2011). Commentary on Valentine, Jeffrey, et al. Replication in Prevention Science. The Advisory Board of Bluepreings for Violence Prevention. Prevetion Science, vol 12, 121-122.

Baekeland, F., & Lundwall, L. (1975). Dropping out of treatment: A critical review. Psychological Bulletin, 82, 738–783.

Chitiyo, M., May, M., Chitiyo, G. (2012). An Assessment of the Evidence-Base for School-Wide Positive Behavior Support. Education and Treatment of Children, Vol. 35, No.1.

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Barrett, S., Bradshaw, C., & Lewis-Palmer, T. (2008). Maryland state-wide PBIS initiative. *Journal of Positive Behavior Interventions, 10,* 1005-114.

Blonigen, B., Harbaugh, W., Singell, L., Horner, R.H., Irvin, L., & Smolkowski, K. 2008). Application of economic analysis to school-wide positive behavior support programs. *Journal of Positive Behavior Interventions*

Bradshaw, C., Reinke, W., Brown, L., Bevans, K., & Leaf, P. (2008). Implementation of school-wide positive behavioral interventions and supports (PBIS) in elementary schools: Observations from a randomized trial. *Education and Treatment of Children, 31,* 1-26.

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes: Results from a

randomized controlled effectiveness trial in elementary schools. Journal Of Positive Behavior Interventions, 12(3), 133-148.

Campbell, D. (1986). Advances in Quasi -Experimental Design and Analysis. New Directions for Program Evaluation, no.31. San Francisco:Jossey-Bass.

Child Trends (2010). Individualized Education Plans. Retrieved from www.childtrendsdatabank.org/?q=node/95 Last update: August, 2010

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155-159.

Cornelius, J. R., Douaihy, A., Bukstein, O. G., Daley, D. C., Wood, S. D., Kelly, T. M., & Salloum, I. M. (2011). Evaluation of cognitive behavioral therapy/motivational enhancement therapy (CBT/MET) in a treatment trial of comorbid MDD/AUD adolescents. Addictive Behaviors, 36(8), 843-848.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. Journal Of Applied Psychology, 78(1), 98-104.

Dunlap, K., Goodman, S., McEvoy, C., & Paris, F. (2010). School-wide positive behavioral interventions and supports: Implementation guide. *Lansing, MI: Michigan Department of Education.*

Flay, B., Biglan, A., Boruch, R., Castro, F., Gottfredson, D., Kellam, S., Moscicki, E.,

Schinke, S., Valentine, J., & Ji, P. (2005) Standards of evidence: Criteria for efficacy, Effectiveness, and Dissemnation. *Prevention Science, vol (10).*

Frey, A., Cloud, R., Lee, J., Small, J., Seeley, J., Feil, E., Walker, H., Golly, A. (2011). The Promise of Motivational Interviewing in School Mental Health. School Mental Health. vol (3).

Green, J., McLaughlin, K. A., Alegría, M., Costello, E., Gruber, M. J., Hoagwood, K., & ... Kessler, R. C. (2013). School mental health resources and adolescent mental health service use. Journal Of The American Academy Of Child & Adolescent Psychiatry, 52(5), 501-510.

Greenspoon, P. J., & Saklofske, D. H. (1998). Confirmatory factor analysis of the multidimensional Students' Life Satisfaction Scale. Personality And Individual Differences, 25(5), 965-971.

Grossman, J., Chan, C., Schwartz, S., Rhodes, J. (2012). The Test of Time in School-Based Mentoring: The Role of Relationship Duration and Re-Matching on Academic Outcomes. American Journal of Community Psychology Vol (49) 43–54

Herrera, C., Grossman, J., Kauh, T. J., & McMaken, J. (2011). Mentoring in schools: An impact study of Big Brothers Big Sisters school- based mentoring. Child Development, 82(1), 346-361.

Huebner. E. S. (1991). Initial development of the student|s life satisfaction scale. School Psychology International, 120-139.

Horner, R., Sugai, G., Smolkowski, K., Todd, A., Nakasato, J., & Esperanza, J., (2009). A Randomized Control Trial of School-wide Positive Behavior Support in Elementary Schools. *Journal of Positive Behavior Interventions.*

Kazdin, A. E. (2011). Evidence-based treatment research: Advances, limitations, and next steps. American Psychologist, 66(8), 685-698.

Karcher, M. J. (2003). The Hemingway: Measure of adolescent connectedness:

Validation studies. ERIC no. ED477969;ERIC/ CASS no. CG032433. Retrieved June 15th, 2009 from http://www.adoelscentconnectedness.com

Kertes, A., Westra, H. A., Angus, L., & Marcus, M. (2011). The impact of motivational interviewing on client experiences of cognitive behavioral therapy for generalized anxiety disorder. Cognitive And Behavioral Practice, 18(1), 55-69.

Kloos, B., Hill, J.L., Thomas, E., & Wandersman, A. (2012). Community Psychology: Linking individuals and communities, 3rd Edition. Menlo Park, CA: Cengage Publishing.

Leary-Tevyaw, T., Monti, P. (2004). Motivational enhancement and other brief interventions for adolescent substance abuse: foundations, applications and evaluation. *Addiction 99. 63-75.*

McFall, R. M. (1991). Manifesto for Science of Clinical Psychology. *The Clinical Psychologist*, *44*, 75-88

Mercer, S.W., Watt, G., Maxwell, M., & Heaney, D. (2004), The consultation and

relational empathy (CARE) measure: preliminary validity and reliability of an empathy-based consultation process measure. *Family Practice*, 21, 699-705.

Merikangas, K., He, J., Burstein, M., Swanson, S., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., Swendsen, J. (2010). Lifetime Prevalence of Mental Disorders in U.S. Adolescents: Results from the National Comorbidity Survey Replication– Adolescent Supplement (NCS-A). Journal of the American Academy of Child & Adolescent Psychiatry. Vol 49 (10).

Miller, L. M., Southam-Gerow, M. A., & Allin, R. r. (2008). Who stays in treatment? Child and family predictors of youth client retention in a public mental health agency. Child & Youth Care Forum, 37(4), 153-170.

Miller, W., Rollnick, S. (2012). *Motivational Interviewing: Helping People Change* (3rd ed.) New York: The Guilford Press.

Mueller, M., Phelps, E., Bowers, E., Agans, J., Urban, J., Lerner, R. (2011).Youth development program participation and intentional self-regulation skills: Contextual and individual

bases of pathways to positive youth development. Journal of Adolescence. Vol. 34. 1115–1125.

McQuillin, S., Smith, B., Strait, G. (2011). Randomized Evaluation of a Single Semester Transitional Mentoring Program for First Year Middle School Students: A Cautionary Result For Brief, School-based Mentoring Programs. Journal of Community Psychology Vol. 39, No. 7, 844–859.

Naar-King, S., Suarez, M. (2011). Motivational Interviewing with Adolescents and Young Adults. New York: The Guilford Press.

National Registry of Evidence-Based Programs and Practices. Retrieved on 9/5/2012 http://www.nrepp.samhsa.gov/

Pajares, F. & Miller, D.M. (1995). Mathematics self-efficacy and mathematics performances: the need for specificity of assessment. *Journal of Counseling Psychology. 42(2)*, 190-198.

Peugh, J. L. (2010). A practical guide to multilevel modeling. Journal Of School Psychology, 48(1), 85-112.

Podell, J. L., Mychailyszyn, M., Edmunds, J., Puleo, C. M., & Kendall, P. C. (2010). The Coping Cat Program for anxious youth: The FEAR plan comes to life. Cognitive And Behavioral Practice, 17(2), 132-141.

Prinz, R., Sanders, M. (2007). Adopting a population-level approach to parenting and family support interventions. *Clinical Psychology Review*, 27, 739–749

Randolph, K., Johnson, J. (2008). School-based Mentoring Programs: A Review of the Research. Children and Schools Vol (30) Number 3.

Rotheram-Borus, M., Swendeman, D., & Chorpita, B. F. (2012). Disruptive innovations for designing and diffusing evidence-based interventions. American Psychologist, 67(6), 463-476.

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *2*(4), 419-436.

Smith, B., Molina, B., Massetti, G., Waschbush, D., Pelham, W. (2007). School-Wide Interventions-The Foundation of a Public Health Approach to School-Based Mental Health. Advances in School-Based Mental Health Interventions: Best Practices and Program Models. Volume 2. Civic Research Institute

Strait, G., Smith, B., McQuillin, S., Terry, J., Swan, S., Malone, P. (2012). A Randomized Trial of Motivational Interviewing to Improve Middle School Students' Academic Performance. *Journal Of Community Psychology*, *40*(8), 1032-1039.

Sugai, G., & Horner, R. (2010). Schoolwide positive behavior supports: Establishing a continuum of evidence-based practices. Journal Of Evidence-Based Practices For Schools, 11(1), 62-83.

Terry, J., Strait, G., Smith, B., McQuillin, S. (In press). Replication of Motivational Interviewing to Improve Middle School Students' Academic Performance. *Journal of Community Psychology. Excepted & Awaiting Publication.*

Urbaniak, G. C., & Plous, S. (2011). Research Randomizer (Version 3.0) [Computer software]. Retrieved on April 22, 2011, from http://www.randomizer.org/

U.S Census School Enrollment Data. http://www.census.gov/hhes/school/ retrieved August 24th 2012.

U.S. Department of Education: http://idea.ed.gov retrieve August 24th 2012.

U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs, *28th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, 2006,* vol. 1, Washington, D.C., 2009.

U.S. Public Health Service, Report of the Surgeon General's Conference on Children's Mental Health: A National Action Agenda. Washington, DC: Department of Health and Human Services, 2000.

Valentine, J., Biglan, A., Boruch, R., Gonzalez, F., Collins, L., Flay, B., Kellam, S., Moscicki, E., Schike, S. (2011). Replication in Prevention Science. *Prevention Science* 12, 103-117.

Walker, H. M., & Shinn, M. R. (2002). Structuring School-Based Interventions to Achieve Integrated Primary, Secondary, and Tertiary Prevention Goals for Safe and Effective Schools. In M. R. Shinn, H. M. Walker, G. Stoner (Eds.) , Interventions for academic and behavior problems II: Preventive and remedial approaches (pp. 1-25). Washington, DC US: National Association of School Psychologist.

Waxman, R. P., Weist, M. D., & Benson, D. M. (1999). Toward collaboration in the growing education-mental health interface. Clinical Psychology Review, 19(2), 239-253.

West, S. G., & Thoemmes, F. (2010). Campbell's and Rubin's perspectives on causal inference. *Psychological Methods*, *15*(1), 18-37.

Weist, M. D., Goldstein, J., Evans, S. W., Lever, N. A., Axelrod, J., Schreters, R., & Pruitt, D. (2003). Funding a full continuum of mental health promotion and intervention programs in the school. Journal Of Adolescent Health, 32(Suppl6), 70-78.

Weist, M. D. (2003). Challenges and opportunities in moving toward a public health approach in school mental health. Journal Of School Psychology, 41(1), 77-82.

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. American Psychologist, 54(8), 594-604.