UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mentor-Based Interventions for Internalizing Problems in Schools: A Research Synthesis

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Arts

in

Education

by

Daniel Kwak

September 2017

Thesis Committee:
Dr. Austin Johnson, Chairperson
Dr. William P. Erchul
Dr. Cathleen Geraghty

The Thesis of Daniel Kwak is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ABSTRACT OF THE THESIS

Mentor-Based Interventions for Internalizing Problems in Schools: A Research Synthesis

by

Daniel Kwak

Master of Arts, Graduate Program in Education
University of California, Riverside, September 2017
Dr. Austin Johnson, Chairperson

The purpose of this research synthesis was to summarize outcomes of mentor-based interventions for children with internalizing problems represented in single-case design studies. Within the school setting, mentors were defined as interventionists who provide directive as well as nondirective support for the students (Barrera & Bonds, 2013). Internalizing problems are described as symptoms of anxiety, depression, and social withdrawal that cause covert forms of emotional distress (Eisenberg et al., 2001; Forns, Abad, & Kirchner, 2011; Levitt & Merrell, 2009). The overall effect (Tau-U = 0.822) across studies demonstrated that mentor-based interventions reduced identified internalizing problems. The moderator variables grade, disability, race, method for student selection, mentor type, mentor's mental health background, frequency of meeting, length of intervention, treatment fidelity, and variations of intervention were analyzed. General findings, moderator analyses, limitations, and implications were discussed.

Keywords: internalizing problems, mentor interventions, single case

Table of Contents

Mentor-Based Interventions for Internalizing Problems in Schools: A Research Synthesis

Problem behaviors can be broadly categorized into externalizing and internalizing behaviors (Achenbach & Rescorla, 2001). Externalizing behaviors are those that are shown in an outward and observable manner such as aggressive and disruptive behaviors (Gresham & Kern, 2004; Hinshaw, 1992). Internalizing behaviors are directed inward, which means problems are experienced within individuals (Cicchetti & Toth, 1992; Forns, Abad, & Kirchner, 2011; Miller & Nickerson, 2007; Reynolds, 1992). Although both internalizing and externalizing behaviors involve characteristics that include behavioral, affective, and cognitive components, externalizing behaviors can be seen as those that may more likely hurt others or disrupt the surroundings, while internalizing problems are those that may be more "intropunitive," which means that the suffering is inflicted on oneself (Zahn-Waxler, Klimes-Dougan, & Slattery, 2000, p. 443).

Internalizing Problems

Definition and characteristics. Internalizing problems can be described as symptoms of anxiety, depression, and social withdrawal that cause covert forms of emotional distress (Eisenberg et al., 2001; Forns, Abad, & Kirchner, 2011; Levitt & Merrell, 2009). Some of these symptoms include depressed mood, feelings of worthlessness, decrease of interest in activities, and fear of social situations (American Psychiatric Association, 2013). However, like any symptomology, internalizing problems are not equivalent to a clinical diagnosis of mental disorders because individuals can experience these symptoms without meeting the distinct criteria and classification established in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5;

American Psychiatric Association, 2013). They can just be characteristics of anxiety and depressive disorders or problems that are not specific to a disorder such as poor self-esteem and negative self-thoughts (Merrell & Gueldner, 2010). Because many of the symptoms may not be observable to educators in schools, the perceptions of students with these problems can be inaccurate.

Perception and nature of the problem. In schools, internalizing problems are sometimes considered to be less important and less problematic in comparison to externalizing problems (Walker & Severson, 1994; Walker et al., 1988). The nature of these types of problems could be a contributing factor of this perception. By nature, internalizing problems are more difficult to detect compared to externalizing behaviors because they are directed inward and not as obviously shown externally (Hunter, Chenier, & Gresham, 2014). Because students with internalizing behaviors are less likely to disrupt class and therefore less likely to interfere with teachers' expectations, the problems the students face are not seen as problematic (Gresham & Kern, 2004; Merrell & Gueldner, 2010). In fact, some traits of these students' behaviors may be considered "good" behaviors because these students may be relatively more quiet, compliant with behavioral expectations, and less challenging to teachers (Gresham & Kern, 2004; Merrell & Gueldner, 2010; Winett & Wrinkler, 1972). Another example would be that the students are simply viewed as shy (Reynolds, 1992). Due to these factors, many students are not referred and thus underserved in schools (Bradshaw, Buckley, & Ialongo, 2008; Kauffman, 1999). This tendency to overlook internalizing behavior problems is problematic due to the prevalence rates and connections to various maladaptive outcomes.

Prevalence. Prevalence estimates are consistent across many studies with approximately 20% of school-age children and adolescents experiencing internalizing problems that significantly impair behavioral, social, or academic functioning (Burns et al., 1995; Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Walker, Nishioka, Zeller, Severson, & Feil, 2000; World Health Organization, 2004). Furthermore, about half of school-age individuals who are diagnosed with a mental disorder suffer from an internalizing disorder specifically with a lifetime prevalence rate of approximately 35% to 50% (Costello et al., 2003; Levitt & Merrell, 2009; Merikangas et al., 2010). By 2020, internalizing disorders are predicted to be the leading cause of illness for school-age children (World Health Organization, 2012). The high prevalence rates and expected increase in these numbers make preventative and remediation efforts important.

Importance in prevention and remediation. Internalizing problems need to be appropriately addressed early on to avoid immediate and long-term consequences because these problems often start in elementary school and continue to escalate if untreated (Greenberg, Domitrovich, & Bumbarger, 2001; Kessler, Berglund, Demler, Jin, & Walters, 2005; Kovacs & Devlin, 1998).

Some immediate consequences in the classroom can be influenced by specific symptoms of internalizing problems. For example, fatigue, irritability, and sadness can impede learning in school by inhibiting students from concentrating in class and reducing participation and engagement, which can ultimately lower test scores and grades (Humensky, Kuwabara, Fogel, Wells, Goodwin, & Van Voorhees, 2010; Linnenbrink, 2006; Schwartz, Rhodes, & Herrera, 2012). Also, symptoms of withdrawal and avoidance

can lead to students' reduced engagement with social activities, decreased interaction with fellow classmates, and reduced attendance (Levitt & Merrell, 2009; Richards & Hadwin, 2011).

These behaviors also have links to other types of problems and possible long-term consequences. First, these "covert" problems could indicate future risk for externalizing behaviors later in childrens' lives (Masten et al., 2005; Kerr, Tremblay, Pagani & Vitaro, 1997). Also, when children suffer from internalizing problems early in their school years, there is increased risk for academic problems, school adjustment, and interpersonal relationships (Beesdo, Bittner, Pine, Stein, Lieb & Wittchen, 2007; Grills & Ollendick, 2002; Hops, Finch, & McConnell, 1985; Horn & Packard, 1985; Muroff & Ross, 2011). They are also linked to reduced school attendance, dropout before completion of high school, substance abuse, risky sexual behaviors, and attempts of suicide (Ingram & Smith, 2008; Kelder et al., 2001; Levitt & Merrell, 2009). If internalizing problems are left untreated, they can affect various components of students' lives and be chronic with possibility for recurrence even if problems subside without treatment (Last, Perrin, Hersen, & Kazdin, 1996). On the other hand, improving student's problems with internalizing issues has been shown to promote learning, prevent onset of problems, and lead to emotional and educational benefits (Atkins, Frazier, & Talbott, 2003; Rones & Hoagwood, 2001). Therefore, it is important to address these problems and schools are arguably the best setting to help these students.

Interventions for Internalizing Problems

Addressing internalizing problems in schools is important because the majority of school-age individuals who receive services for these problems do so in the school setting (Burns et al., 1995; Farmer, Burns, Phillips, Angold, & Costello, 2003; Levitt et al., 2009; Rones & Hoagwood, 2000).

In the schools, various approaches to therapy are implemented to remediate problems with internalizing behaviors. Cognitive behavioral therapy (CBT) is a psychosocial intervention that is widely used in the schools. Examining studies published between 1985 and 1999, Rones and Hoagwood (2000) suggested that the most common way to treat internalizing problems involved implementation of CBT or some form of intervention using cognitive-behavioral strategies.

Origins of CBT. The origins of this approach date back to the 1960s when cognitive theory began to be incorporated in therapy (Mahoney, 1984). By the 1970s, "cognitivism" became pervasive in both research and clinical fields (Mahoney, 1974; Mahoney & Lyddon, 1988; Wilson, 1982). Mahoney and Lyddon (1988) even stated that the field and psychological services underwent a "cognitive revolution" during the 1970s that impacted psychological counseling (p. 191). Internal thoughts and maladaptive cognitive processes were now seen as malleable sources that could be changed (Meyers & Craighead, 1984). For instance, depression was originally thought of as an affective disorder and thinking impairment was seen to be more of a symptom that results from affective disturbance (APA, 1952). Now, there is an increased focus on the specific thought processes and recognition that they are important sources for treatment.

With the rise of cognitivism, behaviorism combined with cognitivism to form the cognitive-behavioral framework. Originally, behaviorism first emerged in the 1910s and developed when observable and objective characteristics were emphasized in the clinical domain and therapy sessions (Hull, 1943; Mahoney & Lyddon, 1988; Skinner, 1963; Watson, 1930). Cognitive as well as cognitive-behavioral approaches became prevalent and different types of therapy were introduced such as Beck's (1970) cognitive therapy, Ellis's (1962) rational emotive therapy, and cognitive behavior modification (Meichenbaum, 1977).

With the development of the cognitive-behavioral approach, the incorporation of cognition into the original behavioral model allowed for broader and effective strategies. According to Kendall and Hollon (1979), CBT was formed to integrate the "efficiencies and methodological rigor of behavioral procedures with cognitive-mediational processes that influence adjustment" (as cited in Courtney et al., 2011, p. 4). Eventually, the continued success of cognitive therapy for adults transitioned to the treatment of children (Courtney et al., 2011).

Characteristics of CBT. With the continued accumulation of evidence since its first emergence, CBT is considered an empirically supported counseling approach that focuses on altering patterns of thoughts and beliefs that are maladaptive, while teaching and reinforcing new, more adaptive behaviors (Barrett et al., 2008; David-Ferdon & Kaslow, 2008; Plotts & Lasser, 2013; Silverman, Pina, & Viswesvaran, 2008). The main principle of CBT is based on the idea that our thoughts create feelings, our feelings create behaviors, and our behaviors reinforce thoughts (Plotts & Lasser, 2013). Because

thoughts and behaviors are considered malleable within the CBT framework (Friedberg & McClure, 2002), efforts are made during sessions to shift them in positive ways. Some of the specific techniques that are implemented include decatastrophizing, test of evidence, disconfirming false beliefs, reattribution, restructuring, problem-solving, and building coping skills (Courtney et al., 2011; Friedberg & McClure, 2002; Greenberg, Domitrovich, & Bumbarger, 2001).

Evidence for CBT. Cognitive behavioral therapy (CBT) has been shown to be effective in addressing childhood and adolescent depression and anxiety in addition to eating disorders, phobic disorders, trauma and anger problems (Cohen, Mannarino, & Deblinger, 2006; Kendall & Hedtke, 2006; Plotts & Lasser, 2013). Comptons, Burns, Egger, and Robertson (2002) found that short-term cognitive behavioral interventions within schools reduced symptoms of depression among children compared to a control group.

Ishikawa, Okajima, Matsuoka, and Sakano (2007) conducted a meta-analysis to examine the effects of CBT for children and adolescents by evaluating the diagnostic interview, severity of symptoms, self-report, parent-report, and teacher-report. The authors found that CBT is an effective treatment for children and adolescents with anxiety disorders as it helped many participants' anxiety to be no longer be clinically significant. Additionally, results from this study suggest that CBT may be effective for children who have comorbid depressive symptoms. Spielmans, Pasak, and McFall (2006) also conducted a meta-analysis and found that CBT as well as other bona-fide therapies (e.g., social effectiveness therapy, interpersonal therapy, and systemic behavior family

therapy) are effective interventions for children and adolescents in addressing depression and anxiety when compared to non-theory based interventions. In this context, therapies were considered bona fide if they met criteria established by Wampold et al. (1997). Some requirements were that the therapist had to be trained to provide the therapy, provided individualized treatment through face-to-face meetings, and delivered treatments that contained psychologically valid components.

Knowing that CBT is an effective intervention for youth, the question is whether there is a need for a different intervention. Specifically, it is appropriate to discuss why interventions that include a mentor component can be beneficial for students suffering from internalizing concerns.

Benefits of Mentor Interventions in Comparison to Counseling

The current literature on interventions targeting internalizing problems is largely limited to interventions that involve some form of one-on-one or group counseling. Although mentor-type interventions do not currently have as much evidence for treating internalizing behaviors as opposed to externalizing behaviors, elements of this intervention that differ from counseling type interventions may make them more feasible and resourceful in addressing internalizing problems.

Mentor-type interventions are more flexible in terms of who can serve as mentors, as they do not require a high degree of expertise. For school personnel who are not school psychologists or school counselors, accurate implementation of CBT may be difficult (Maag & Swearer, 2005). Therefore, mentor-type interventions have the potential to utilize interventionists who may not have advanced knowledge of mental health. This

could mean that more students suffering from internalizing problems may be served if educators in schools are open to taking part in these interventions.

Additionally, counseling-type interventions may take longer and cost students valuable instruction time. Cognitive-behavioral interventions or counseling-type interventions generally tend to last more than 15 sessions and can take an hour or more per week during class time (Ginsburg & Drake, 2002; Merrell & Gueldner, 2010). Although the amount of time spent weekly with mentor-type interventions may not be significantly less than with counseling-type interventions, some studies have shown that positive effects can be seen in as little as a few weeks (Cook et al., 2015; Herrera & Karcher, 2013; Hunter et al., 2014). Although each counseling and mentor-type interventions varies in time and length, it seems that mentor-type interventions may take slightly less time to see an effect. Also, counseling interventions require students to leave class to receive intervention, and so students miss instructional time on a regular basis. Although mentor-type interventions can also be conducted in this format, some formats like the ones used in Check in Check Out (Hunter et al., 2014), can prevent missed class time. These benefits of mentor-type interventions make them possible alternative or additional treatment to counseling interventions.

Definition of a Mentor in the School Setting

There is a wide variety of proposed definitions that describe a mentor. Rhodes (2002) defines a mentor as an adult who "provides ongoing guidance, instruction, and encouragement aimed at developing the competence and character of the protégé" (p. 3). The mentor would also "[serve] as a role model and advocate" for the mentee (Rhodes,

2002, p. 35). Hoyle, Marshall, and Yell (2011) explained that mentoring relationships provide social and academic support for the mentee. Some other definitions incorporate specific suggestions for difference in age between the mentor and mentee, acceptability of compensation by the mentor, and duration of the mentor-mentee relationship (Dubois & Karcher, 2013).

Despite the differences, there are "recurring themes" or core elements of mentoring that can be found in many of the proposed definitions. One commonly agreed on idea is that mentors need to have more knowledge or experience compared to the mentee (Dubois & Karcher, 2013). This would allow the mentor to offer guidance that will result in the development of the mentee (Dubois & Karcher, 2013). Also, a mentor-mentee relationship may be characterized by the existence of an emotional bond through social support (Dubois & Karcher, 2013). To develop this relationship, the engagement of specific activities may vary widely from teaching discrete skills to watching a movie together on an outing.

Typical activities of mentors in the community setting may not be appropriate or feasible in the school setting. For instance, school-based mentors would most likely not engage in an outing with a mentee on a weekend. Therefore, within the school context, the duties of a mentor may differ substantially. In the school setting, essential forms of support would be likely be provided in a more structured manner and involve regular meetings (Portwood & Ayers, 2013). Additionally, in schools, the duration of the mentor-mentee relationship may be significantly shorter because the relationship, seen as an intervention, will likely end once target goals are reached (Portwood & Ayers, 2013).

The purpose of this study necessitates an explicit definition of a mentor within the school setting. As mentioned earlier, mentors should preferably have greater knowledge or experience generally, or in the specific area mentees need support in. Also, Barrera and Bonds (2013) suggested that mentoring would involve directive as well as nondirective guidance. The authors stated that directive guidance is used to improve certain skills of the mentee, while nondirective support is used to establish rapport or build emotional connection. Taking this into consideration, mentors would need to provide directive support, defined as explicit teaching or training of skills, as well as nondirective support for the purposes of this study. To address the nondirective component, mentees would need to have the opportunity to converse or interact with the mentor without structured time set aside just for explicit training or teaching of skills (e.g., conversing about the day). The purpose of having both components is to build some feelings of social support and connection, while still learning from the mentor and the intervention.

Mentor Interventions

Although there is evidence to suggest that mentor interventions lead to positive outcomes, the intervention effects have been reported as small or moderate when evaluated according to Cohen's (1988) guidelines. Tolan, Henry, Schoeny, Lovegrove, and Nichols (2013) conducted a meta-analysis to examine the effects of mentoring interventions on delinquency as well as other related problems such as aggression, drug use, and academic achievement. In this analysis, the target students were specifically those at risk for or currently involved in delinquent behavior. The independent variable

was an intervention that at the very least included mentoring as one component of the intervention, but not those that focused on psychotherapy, cognitive-behavioral therapy, or behavior modification. The dependent variable was one of the following: delinquency, aggression, substance use, or academic achievement. From the analysis of 46 articles included in the study, the authors concluded that mentoring can have significant positive effects for high-risk youth in reducing the problems identified, but the effects were moderate for all four categories. Of the four categories, the intervention yielded the highest effect sizes for aggression, followed by delinquency, drug use, then academic achievement outcome measures.

Dubois, Holloway, Valentine, and Cooper (2002) also conducted a meta-analysis to examine the effect of mentoring for youth. In this analysis, the only specific criterion established for target students was that the mean age of the students in each study had to be less than 19. The independent variable was a one-on-one mentoring intervention with the exclusion of interventions that were mediated by peers as opposed to older youth or adults. The dependent variable was one of five outcome measures including "emotional/psychological, problem/high risk behavior, social competence, academic/educational and career/employment" type outcomes (Dubois et al., 2002, p. 183). From the analysis of 55 articles included in the study, the results suggested that the interventions yielded small effects. Out of the five categories, the intervention yielded highest effect sizes for career/employment, followed by problem/high-risk behavior, social competence, academic/educational, then emotional/psychological outcome measures. Overall, there is a relatively limited number of studies that specifically address

internalizing problems within or outside the context of schools (Dubois et al., 2002; Tolan et al., 2013).

School-based Mentor Interventions

The definition developed for the purpose of this study fit the features of some established mentor-type interventions like Check In Check Out (CICO) and Check and Connect (Hoyle et al., 2011). Variations of CICO have been established as effective interventions for addressing externalizing behaviors, and the discussion of these interventions is important because the majority of the studies used in this research synthesis represent some variation of CICO.

Check In Check Out (CICO) is used primarily as a Tier 2 intervention in the context of positive behavior supports (Crone, Horner, & Hawken, 2010). The purpose of this intervention is to "improve the overall efficiency of the school-wide procedures, while reducing the number of individualized interventions that are needed" (Crone et al., 2010, p.1). Many published articles provide support for the effectiveness of CICO interventions (Hawken, 2006; Hawken & Horner, 2003; Hawken, MacLeod, & Rawlings, 2007; March & Horner, 2002; Todd et al., 2007).

CICO is an intervention that provides daily support and monitoring to students (Crone et al., 2010). There are basic features and procedures that make up a standard CICO intervention. First, students "check in" with a facilitator before classes begin (Myers & Briere, 2010). The interventionist reviews expectations, sets performance goals, and gives the student a point sheet to carry to gather performance feedback from teachers throughout the school day (Myers & Briere, 2010). At the end of the day, students

"check out" with the interventionist (Myers & Briere, 2010). This process may involve the interventionist reviewing the point sheet for the day, acknowledging accomplishments, and providing students' preferred reinforcers (e.g., activities, privileges, or tangible items; Myers & Briere, 2010). CICO interventions are typically implemented 3-5 days per week, and may only require 5-10 minutes from the interventionist (Crone et al., 2010). Crone and colleagues (2010) explained that CICO interventionists may be able to support around 20 students.

Although multiple evaluations of mentor-based interventions have been published (Dubois, Holloway, Valentine, & Cooper, 2002; Hawken, Bundock, Kladis, O'Keeffe, & Barrett, 2014; Maggin, Zurheide, Pickett, & Baillie, 2015; Tolan, Henry, Schoeny, Lovegrove, & Nichols, 2013; Wolfe, Pyle, Charlton, & Sabey, 2016), sufficient evidence has only been established for remediating externalizing behaviors. Therefore, one of the goals of this research synthesis is to examine the effectiveness of mentor-based interventions in helping students with internalizing problems.

Research Questions

The purpose of this research synthesis is to evaluate the following:

1. To what extent are mentor-based interventions for children with internalizing problems in schools, represented in single-case experimental research, effective?
2. To what extent does their effectiveness differ by student grade, student diagnosis, student race, method for student selection, mentor type, mentor's background in mental health, frequency of meeting, length of intervention, and fidelity of treatment?

Method

Search Procedure

To determine the articles to be included in the research synthesis, Educational Resources Information Center (ERIC) and PsycINFO online databases were systematically searched for peer-reviewed studies. A limitation to the search period was not made because it was expected that there would be a limited number of mentor-based intervention studies that address internalizing problems in the school setting. The keywords used to search the two databases were "internalizing problem," "internalizing behavior," "internalizing disorder," and "socially withdrawn." These keywords were paired with the following terms: "single case," "mentor program," and "check in check out." This search yielded a total of approximately 7,000 abstracts. Each of these abstracts was examined to determine if the study implemented a mentor-based intervention, addressed internalizing problems, and used a single-case design. If this information was not clear from reading the abstracts, the full article was reviewed. The selected articles were reviewed in detail to determine that they meet the inclusion criteria indicated below. References of the selected articles were also reviewed to search for additional articles.

Inclusion Criteria

To be included in the research synthesis, studies had to meet eight criteria. The first three criteria are taken directly from the standards established by the What Works Clearinghouse for single case designs (Kratochwill et al., 2010).

1. The study used a single-case design and included "at least three attempts to demonstrate an intervention effect at three different points in time or with three different phase repetitions" (Kratochwill et al., 2010, p. 15).
2. "The independent variable (i.e., the intervention) must be systematically manipulated, with the researcher determining when and how the independent variable conditions change" (Kratochwill et al., 2010, p. 14).
3. "For a phase to qualify as an attempt to demonstrate an effect, the phase must have a minimum of three data point." (Kratochwill et al., 2010, p. 15).
4. The study provided all baseline and treatment data points to calculate Tau-U.
5. The study included a mentor-based intervention. To be considered as a mentor-based intervention, mentors needed to provide directive (e.g., explicit teaching or training of skills) as well as nondirective support (e.g., conversing about the day).
6. The intervention took place in a public, non-residential school setting.
7. The intervention targeted at least one identified internalizing problem.
8. The study reported data for at least one dependent measure that assessed an internalizing problem. As previously mentioned, internalizing problems can be described as symptoms of anxiety, depression, and social withdrawal that cause covert forms of emotional distress (Eisenberg et al., 2001; Forns, Abad, & Kirchner, 2011; Levitt & Merrell, 2009).

A total of 8 studies met these criteria (Table 1).

Interobserver agreement was not incorporated as a criterion for inclusion. One reason for this decision was that not all studies addressing internalizing problems were

predicted to have outcome measures that may be observable by nature. Also, because a limited number of articles was expected to be found, further restrictions were avoided in order to be more inclusive. A standard for the reliability of instruments used to measure the outcome variables was not incorporated for the same reason. Additionally, it was considered that having stricter criteria and limiting the number of studies could possibly lead to loss of data from excluded studies that could potentially be useful (Lipsey & Wilson, 2001). In these cases, it is important to examine the extent to which results differ based on the existence or absence of interobserver reliability data (Lipsey & Wilson, 2001). Sometimes, if the results of groups of studies differ considerably, it could be beneficial to focus on the results of the studies that are considered better in quality. However, it is worth noting that the outcomes (i.e., effect sizes) for the studies with and without interobserver reliability were similar (Table 1). Even so, this is a significant limitation and further discussion on interobserver agreement can be found in the limitations section.

Dependent and Independent Variables

The dependent variables measured in the studies all dealt with the remediation of specific internalizing problems, but were varied in their approach. There were three types of dependent variables, which included measuring rates of appropriate behavior (e.g., effective communication, socially appropriate behavior, and positive social engagement), points on a report card made up of a number of goals (e.g., Daily Progress Report, Daily Behavior Report, and Daily Behavior Report Card), and self-report rating on the students' level of distress (e.g., Subjective Units of Distress). With the report cards

such as daily progress report (DPR), multiple goals were able to be combined to generate one score reflecting students' overall performance through the use of the point system. For instance, if a student met four out of five goals, his or her score would be 4 or 80%. This way, there would still be just one value representing one dependent variable. Some of the specific goals found in these report cards include asking for hcelp when needed, starting conversations with peers, and participating appropriately in class (Collins, Gresham, & Dart, 2016; Hunter et al., 2014).

Some studies provided more than one dependent measures, in which case only one was used to calculate the effect size. Two out of the eight studies included in the analysis measured multiple dependent variables over time (Marchant et al., 2007; Ross & Sabey, 2014). Marchant et al. (2007) reported effective communication and appropriate peer play as outcome measures. Between the two, effective communication was selected as the target outcome measure because the operational definition was judged to be clearer and the behavior was considered to be more generalizable. Ross and Sabey (2014) reported positive social engagement and negative social engagement as outcome measures. Positive social engagement was selected as the target outcome measure because the authors stated that it was the primary dependent variable. Furthermore, it was judged that measurement of positive social engagement would take into consideration, to some extent, negative social engagement as well.

The independent variable was mentor-based interventions or interventions that include a mentor component with the mentor providing directive as well as nondirective support to the students. Specific interventions included in the analysis can be divided into

three categories: standard CICO (Collins et al., 2016; Dart et al., 2015), adjusted CICO (Cook et al., 2015; Fiat et al., 2017; Hunter et al., 2014; Ross & Sabey 2015), and social skill-focused interventions (Christensen et al., 2007; Marchant et al., 2007). The standard CICO would follow procedures that are essentially identical to the sample procedure provided previously. Adjusted CICO interventions, on the other hand, either incorporated CBT techniques or added a social skills component to the standard CICO procedure. For example, when a student faced internalizing problems that he or she was not able to resolve the day before, the mentor would work with the student to identify negative thoughts and replace them with positive ones (Hunter et al., 2014). Another adjusted CICO intervention adopted a similar format, but also initially included two 40 minute sessions that incorporate CBT content to start out the intervention (Cook et al., 2015; Fiat et al., 2017). For specific details of independent and dependent variables used in the studies, refer to Tables 3, 4, 5 and 6.

Interrater Agreement

Inclusion criteria and coding of study information were evaluated for interrater agreement. Interrater agreement was determined by dividing the number of agreements by the number of agreements and disagreements. Subsequently, the resulting number was multiplied by 100 to derive a percentage. A graduate student evaluated three (38%) of the eight articles included in the research synthesis for interrater agreement. Interrater agreement was 88% for inclusion criteria and 92% for coding.

Coding and Moderators

Overall, information collected and coded provided information regarding the sample population, mentors, and interventions were collected. Each of the studies was coded on the student's grade, student's diagnosis, student's race, method for student selection, mentor type, mentor's mental health background, frequency of meeting, length of intervention, and treatment fidelity. Because six of the eight studies incorporated some variation of the CICO intervention, it was coded whether or not the CICO intervention was adapted to target internalizing problems. Although not included as a moderating variable, schools' implementation of Positive Behavior Interventions and Supports (PBIS) was also noted.

Each of the moderating variables was divided into two or three levels. The moderating variables are explained below:

1. Student's grade level: This variable was divided into three levels including preschool through third grade, fourth grade through sixth grade, and seventh grade through high school.
2. Student's disability: The two levels of this variable were categorized into students who are identified with a disability and those who are not.
3. Student's race: Students either identified as Caucasian or Minority, which included African Americans, Hispanics, and Asians.
4. Method for student selection: This variable differentiates the method in which the students were identified for inclusion within the study. Identification by nomination or screening was categorized into one category, while identification

by the use of a multiple gating procedure was categorized into another. It is important to note that for the first category noted, nominations or screening involved identification of student through a single-step process.

5. Mentor type: This variable distinguishes whether the mentor was an adult or a peer.
6. Mentor's mental health background: This variable categorizes mentors who most likely have advanced knowledge in mental health and those who likely do not. School psychologists, school counselors, and psychology students were considered to have greater knowledge of mental health compared to teachers and peers.
7. Frequency of meeting: The three levels of this variable were 3 to 5 days per week, 1 to 2 days per week, and less than once per week.
8. Length of intervention: The two levels of this variable were 10 days or less and greater than 10 days.
9. Treatment fidelity: This variable examines the extent to which the intervention was implemented with fidelity. It was divided into two levels. At one level, the mentor's fidelity of treatment would need to meet the minimum criterion of 80%. At the other level, this criterion would not be met.
10. Check In Check Out variation: This variable differentiates CICO interventions that were not modified and those that were modified to address internalizing problems by adding a component to the intervention that specifically address students' thoughts and feelings.

Effect Size Calculation

Tau-U. Tau-U was used to calculate effect sizes for the research synthesis. Because the results were weighted, the range of the effect sizes was between -1 and 1. Tau-U is a nonparametric procedure that examines data nonoverlap between baseline and intervention phases, utilizing all observation data from both phases (Parker, Vannest, Davis, & Sauber, 2011). To sum it up in one phrase, Tau-U yields "non-overlap after controlling for Phase A trend" (Parker, Vannest, & Davis, 2011, p. 314). The distinct characteristic of Tau-U is that it adjusts for the baseline trend and incorporates information about the trend of the intervention phase (Parker et al., 2011). The four different indices used to calculate the effect size include nonoverlap between baseline and intervention phases, nonoverlap and trend of intervention phase, nonoverlap with controlled baseline trend, and nonoverlap and controlled trend of intervention phase (Parker et al., 2011). Tau-U was used to calculate the overall effect across studies, of each individual study, and each level of the moderator variables. The Tau-U weighted effect sizes were calculated through a web-based calculator on www.singlecaseresearch.org, a website developed by Vannest, Parker, and Gonen (2011).

Rationale for using Tau-U. Vannest and Ninci (2015) explained why this method for calculating effect sizes is preferred in single-case research designs. First, the authors state that Tau-U adjusts for trend, which can yield important changes in the effect size. Compared to other effect size calculations, Tau-U is more robust for smaller data sets and better "[discriminates] at the upper and lower limits and [correlates] with other indices" (Vannest & Ninci, 2015, p.407).

Parker et al. (2011) reviewed nine effect-size indices (i.e., ECL, PND, PAND, Phi, PEM, IRD, NAP, Tau_{novlap}, Tau-U) and determined that Tau-U most effectively allows for reliable detection of smaller effects with greater statistical power through Kendall's S test. Tau-U and ECL are the only two methods that control for the baseline phase compared to other nonoverlap methods discussed. However, Tau-U is unique in that it controls for monotonic trend, while ECL is only able to control for linear trends specifically (Parker et al., 2011).

Baseline trend correction. Prior to conducting effect size calculations, baseline trends were analyzed to check if there was a need for baseline correction. Baseline correction is necessary because within single-case designs, the reliability of the baseline phase needs to be determined in order to attribute the trend or change in outcome at the intervention phase to the treatment. For example, if the desired outcome is an upward trend and the baseline phase already has a significant upward trend, it is more likely that that this trend will continue even without intervention. In this case, a correction is made so that the effect size is not an overestimate. Parker and colleagues (2011) noted that the baseline trend should be corrected if the trend is statistically significant. Similarly, Tarlow (2016) explained that if there is a significant trend in the baseline, that trend should be removed. In this process, the residuals of both the baseline and intervention phases are calculated from the regression line of the baseline phase (Tarlow, 2016). Among the students represented in the studies, one student's baseline trend from Collins et al. (2016) was corrected through the web-based calculator that was previously mentioned, and the corrected contrasts were used for effect size calculations.

Results

Study Information

Each of the studies included in the research synthesis were published in different journals. The journals are *Education and Treatment of Children*, *Behavior Modification*, *School Mental Health*, *School Psychology Quarterly*, *Journal of Applied School Psychology*, *Journal of Emotional and Behavioral Disorders*, *Psychology in the Schools*, and *Remedial and Special Education*. There were approximately 4 participants in each study with a minimum of 1 participant and maximum of 6. On average, there were approximately 6 baseline data points and 13 treatment data points.

Study designs. One of the studies (i.e., Christensen, Young, & Marchant, 2007) included in the analysis utilized simple phase change design (i.e., ABAB) with alternations between the baseline and intervention phases. In this design, the effect of the treatment is established multiple times through the implementation (Kratochwill et al., 2010).

The remaining seven studies included in the analysis utilized multiple-baseline designs. In this design, there is a series of baseline and intervention effects conducted at the same time where the start of the intervention phase would begin at successive time points for a minimum of two cases (Kratochwill et al., 2010). The staggering of treatment implementation for each successive participant allows for attribution of change in the outcome to the treatment.

Overall Outcome

A total of 29 outcomes of participants from the 8 selected studies were used for analysis to estimate treatment effects. The overall Tau-U across all selected studies was 0.822 with a 90% confidence interval between 0.726 and 0.918.

Descriptive Characteristics and Moderator Analysis

Baseline condition. Four out of the eight studies did not specifically identify whether or not Positive Behavioral Interventions and Supports (PBIS) was being implemented before the start of interventions, while the other four studies indicated that PBIS was in place.

Participant characteristics. A total of 29 individuals across the studies participated in mentor-based interventions. Gender and race were reported in all 8 studies. Across the studies, most participants were female (59%, $n = 17$). Participants were identified as Caucasian (45%, $n = 13$), African American (28%, $n = 8$), Hispanic (17%, $n = 5$), and Asian (10%, $n = 3$). All studies reported student grade. The distribution of the grades was mostly limited to the elementary level; 24% ($n = 7$) of students ranged from preschool to third grade, 62% ($n = 18$) of students ranged from fourth grade to sixth grade, and 14% ($n = 4$) of students ranged from seventh grade to high school. Some studies reported student disability. All students with a disability were identified as having a learning disability 14% ($n = 4$). Across studies, 48% ($n = 14$) of the students were identified through multiple gating processes, 28% ($n = 8$) were identified through nomination, and 24% ($n = 7$) were identified by a screening score.

Method for student selection, race, grade, and disability analysis. To determine if the results of the intervention varied by method of student selection, Tau-U coefficients for multiple gating and single-step selection (i.e., screening or nomination) were evaluated. Multiple gating method showed a greater effect (Tau-U = 0.857; CI 90% range, 0.726 to 0.987; n = 14) compared to screening or nomination (Tau-U = 0.785; CI 90% range, 0.648 to 0.925; n = 15).

To determine the extent to which race moderated intervention effectiveness, results of Caucasian participants were compared to the other participants (i.e., African American, Hispanic, and Asian). The participants identified as Caucasian showed a greater effect (Tau-U = 0.875; CI 90% range, 0.730 to 1.000; n = 13) compared to other participants categorized as "Minority" (Tau-U = 0.778; CI 90% range, 0.649 to 0.906; n = 16).

Student grade was also examined as a moderator variable. The range of participants in 4th to 6th grade showed the greatest effect (Tau-U = 0.847; CI 90% range, 0.697 to 0.998; n = 18), followed by those in the 7th grade to high school range (Tau-U = 0.795; CI 90% range, 0.508 to 1.000; n = 4), then lastly, those in the preschool to 3rd grade range (Tau-U = 0.776; CI 90% range, 0.557 to 0.995; n = 7).

Student disability was used as a moderator to determine if the indication of a disability of the participants in the studies would yield differential effects for the two groups of participants. Participants who were not identified with a disability showed a greater effect (Tau-U = 0.835; CI 90% range, 0.732 to 0.938; n = 24) compared to those

with a disability, all of whom were identified as having learning disabilities (Tau-U = 0.735; CI 90% range, 0.473 to 0.998; *n* = 4).

Mentor characteristics. Mentor type was reported in all eight studies. Depending on the study, each mentor was either assigned to one or multiple students. Across the studies, 38% (*n* = 8) of the mentors were peers, 19% (*n* = 4) were student teachers, 14% (*n* = 3) were school psychology graduate students, 5% (*n* = 1) was a school psychologist, 5% (*n* = 1) was a counselor, 5% (*n* = 1) was a special education teacher, 5% (*n* = 1) was a paraprofessional, 5% (*n* = 1) was an undergraduate student majoring in psychology, and 5% (*n* = 1) was an undergraduate student majoring in special education.

Mentor type and mentor mental health background analysis. To determine if the type of mentor moderated intervention effects, adult-mediated interventions were compared to peer-mediated interventions. Adult-mediated interventions showed a greater effect (Tau-U = 0.864; CI 90% range, 0.751 to 0.977; *n* = 21) compared to peer-mediated interventions (Tau-U = 0.709; CI 90% range, 0.527 to 0.892; *n* = 8).

Mentor's background in mental health was also examined, and it was determined that having more advanced knowledge in mental health yielded greater intervention effect (Tau-U = 0.864; CI 90% range, 0.751 to 0.977; *n* = 4) compared to the counterpart (Tau-U = 0.709; CI 90% range, 0.527 to 0.892; *n* = 12).

Intervention characteristics. Across the 8 studies, 50% (*n* = 4) of the interventions used modified CICO interventions that address internalizing problems, 25% (*n* = 2) of the interventions used standard CICO intervention, and 25% (*n* = 2) addressed social skills for concerns regarding withdrawal.

Meeting frequency, intervention length, treatment fidelity, and CICO variation analysis. Frequency of meeting with the mentor was examined to determine if it moderated intervention effects. Results indicate that meeting 3-5 days per week has a greater effect (Tau-U = 0.824; CI 90% range, 0.715 to 0.934; n = 22) when compared to meeting 1-2 days per week (Tau-U = 0.813; CI 90% range, 0.611 to 1.000; n = 7). The third level for this variable was not included in the analysis because none of the participants met with their mentor less than once per week.

To determine the extent to which the length of intervention moderated intervention effectiveness, results of different lengths of intervention were evaluated (more than 10 days and 10 days or less). The greater effect was seen with interventions that were carried out for 10 days or less (Tau-U = 0.862; CI 90% range, 0.665 to 1.000; n = 8) compared to those that lasted more than 10 days (Tau-U = 0.808; CI 90% range, 0.698 to 0.917; n = 21). Treatment fidelity was also examined and when there was no fidelity data provided or mentor's fidelity of treatment did not meet a minimum criterion of 80%, it produced greater intervention effects (Tau-U = 0.813; CI 90% range, 0.661 to 0.960; n = 10) compared to when mentor's fidelity of treatment met a minimum criterion of 80% (Tau-U = 0.785; CI 90% range, 0.674 to 0.895; n = 22)**.**

Because the majority of the studies incorporated some variation of the Check In Check Out (CICO) intervention, they were specifically examined to determine if there are differential effects when adjustments were made to address internalizing problems. Results indicate that CICO with modifications yield greater effects (Tau-U = 0.833; CI

90% range, 0.706 to 0.960; *n* = 18) compared to standard CICO interventions (Tau-U = 0.697; CI 90% range, 0.499 to 0.894; *n* = 7).

Outcomes Based on Studies

A range of Tau-U coefficients were found for each of the 8 studies (i.e., 0.658, 0.729, 0.792, 0.685, 0.817, 0.853, 0.881, 0.993). The study by Marchant and colleagues (2007) yielded the highest effect (Tau-U = 0.993; CI 90% range, 0.778 to 1.000; *n* = 3). The study by Dart and colleagues (2015) yielded the lowest effect (Tau-U = 0.658; CI 90% range, 0.379 to 1.000; *n* = 3). There are a few notable differences between these two studies. Marchant et al. (2007) addressed problems with social withdrawal with an adult mentor, while Dart et al. (2015) addressed a range of internalizing problem behaviors with a peer mentor. Additionally, the intervention used in Marchant et al. (2007) included a component where the adult mentor actively engages in training and teaches student to use social skills to interact with others at the playground. On the other hand, Dart et al. (2015) used a CICO procedure that does not make specific adaptions to treat internalizing problems other than the use of goals on the DBR that relate to internalizing behaviors. Therefore, these differences may have attributed to the differences in Tau-U coefficients between the two studies.

Interpretation of Effect Size

Benchmarks for effect sizes are often used for interpretations. For instance, Cohen's (1988) guidelines suggest a specific criterion for evaluating strength of effect sizes when evaluating standardized mean differences—an effect size of 0.20 is considered small, 0.50 medium, and 0.80 large. With the PND method for single-case

research, an effect size may be considered large if it is above .90, moderate if it is between .70 and .90, and small if it is below .70 (Scruggs, Mastropieri, & Casto, 1987, as cited in Vannest & Ninci, 2015). However, simply describing effect sizes in accordance to such criteria is not practical because the significance of the effect may depend heavily on the context of the study. Vannest and Ninci (2015) indicated that instead of using these set criteria, it is more meaningful to directly interpret findings in relation to client needs, context, and prior intervention work.

Therefore, moderator analysis is particularly helpful in determining the relative effectiveness of interventions based on certain characteristics of the student, mentor, or intervention. Because the levels of the moderators are compared with the same effect size calculation and are within the context of the same study, direct comparisons can be made. In evaluating the overall effect, studies examining mentor-based interventions and forms of therapy that address internalizing problems will be considered. Although strength of effect sizes will be noted and guidelines for evaluating effect sizes will be utilized, the overall effect will be analyzed in relation to similar interventions within the appropriate context.

Discussion

Discussion of Overall Effect

The first research question was developed to summarize the effects of interventions utilizing mentors to address internalizing problems in the school setting. Although direct comparisons of Tau-U coefficients within single-case designs cannot be made due to the limited research on this topic and use of this specific analysis (Tau-U),

results of other published studies can be considered to provide context for the effect of the intervention.

To examine mentor interventions, two groups of authors used Cohen's (1988) guidelines to evaluate the findings of their meta-analyses. Tolan and colleagues (2013) found that these interventions yielded mostly small effects for delinquency (*SMD* = 0.21), aggression (*SMD* = 0.29), drug use (*SMD* = 0.16), and academic achievement (*SMD* = 0.11). Similarly, Dubois et al. (2003) reported small to moderate effect sizes (*d* = 0.18) when they examined the effect of mentoring interventions for youth with greater effects observed when specific goals were pursued (*d* = 0.21) compared to more general goals (*d* = 0.14).

With mentor-based interventions in the schools, Hawken et al. (2014) found that the effect sizes were small (*d* = .37) for group studies according to Cohen's (1988) guidelines and moderate (*PND* = .68) for single-case studies according to criterion suggested by Scruggs et al. (1987) in addressing problem behavior, academic engagement, and work completion among other concerns. Additionally, Maggin et al. (2015) found that problem behaviors were reduced with NAP estimate of 0.83, IRD estimate 0.62, and *SMD* estimate of 1.46. According to Wolfe et al. (2016), CICO interventions yielded Tau-U weighted mean coefficients that ranged from 0.30 to 0.89 in single-case designs, while the interventions yielded Cohen's *d* value of 0.40 in group designs for addressing externalizing behaviors.

With counseling interventions, Ishikawa and colleagues (2007) found that CBT had a moderate effect (*d* = 0.77) in treating anxiety disorders when compared to the

control group. Weisz, McCarthy, and Valeri (2006) found that psychotherapy had a small effect for children and adolescents with depression (*ES* = .34). In addition, Kim (2008) conducted a meta-analysis to evaluate the effectiveness of solution-focused therapy and found that it yielded a small effect (*ES* = .26) for internalizing problems.

Overall, mentoring interventions tend to yield relatively small effects, while mentor-based interventions in schools and counseling interventions tend to yield slightly greater effects. Still, all three types of interventions, evaluated through various meta-analyses, seem to show modest effects. In the context of these results, the overall effect of the current study appears to be similarly effective.

Discussion of Moderator Effects

Mentor characteristics. Intervention effects were greater when the interventions were mediated by an adult compared to a peer. Moreover, guidance and supervision were still provided by an adult to the mentor if the designated interventionist was a peer. For instance, a teacher would still be responsible for providing DBRC forms, informing the peer mentor of the goals, and supervising the "check out" process (Collins et al., 2016). Sometimes, an adult also needed to supervise both the check in and check out process (Dart et al., 2014). Therefore, it may be more efficient for an adult to provide direct intervention, since considerable effort may be needed just for supervision of peer mentors.

Mentors' mental health backgrounds were also examined in the analysis. The problem, though, was that some of the studies did not identify specific assignment of mentors to students or indicated that mentor assignments were randomized. Therefore, the number of participants within this moderating analysis was limited. Although school

psychologists and school counselors are considered to have knowledge in mental health, they were not included because the mentor assignments were unclear in studies utilizing these mentors. Within this moderator analysis, school psychology graduate students were considered to have background in mental health, while teachers and peers were considered to have less. This comparison yielded a greater effect size for the first group of mentors as opposed to the latter.

The aforementioned lack of specification was not the only problem encountered in attempting to extract information on characteristics of the mentors. Generally, information reported regarding the mentors were very limited. Three studies (Christensen et al., 2017; Collins et al., 2016; Dart et al., 2014) reported information on mentor's race/ethnicity, four studies (Christensen et al., 2017; Collins et al., 2016; Dart et al., 2014; Marchant et al., 2007) reported information on mentor's gender, and three studies (Collins et al., 2016; Dart et al., 2014; Marchant et al., 2007) reported information on mentor's age. Because these interventions place a heavy emphasis on the mentors, their characteristics may possibly mediate the effects of the intervention. It may be the case that these characteristics are less important in the school setting. For instance, the matching of certain characteristics of the mentor to the student may be not be realistic in the school setting compared to in the community setting. However, the knowledge of the differential effects would still be valuable.

Student characteristics. The interventions were more effective for Caucasian students as compared to students from other racial groups (i.e., African American, Hispanic, and Asian). The intervention may have been more effective for Caucasian

students as a result of factors not identified in the studies. Also, effects were greater for 4th to 6th grade students compared to preschool to 3rd and 7th to high school students. As mentioned previously, the vast majority of the participants were elementary school students.

Intervention-related characteristics. The effectiveness of the intervention was similar when implemented for less than 10 days as opposed to longer durations of time. Although this may be surprising because most of these interventions were implemented daily, the frequency of the intervention may have yielded benefits early on. Also, higher treatment fidelity did not yield a greater effect. This may have been the case because only one study in the "less than 80% treatment fidelity or not collected" category reported treatment fidelity data. It is possible that the studies that did not report this information still implemented the intervention with high treatment fidelity.

Among the studies that used variations of the Check in Check Out intervention, comparisons were made between those that did not incorporate modifications (aside from the goals identified reflecting internalizing problems) to those that did adjust the intervention to specifically address internalizing problems. The results suggest that modified CICO interventions were more effective compared to those that were not modified. An example of modification to the intervention is identifying maladaptive thinking patterns and working to promote problem-solving strategies during the check-in process when necessary (Hunter et al., 2014). Another example is to initially implement two treatment sessions to provide psychoeducation (Cook et al., 2012). Students would learn to normalize emotions and understand their emotions by using coping skills. This

would be followed by daily implementation of the standard CICO procedure with the option to practice coping skills when needed.

The selection methods of the students to be included in each of the studies were also analyzed. The results suggest that those who were identified through multiple gating selection gained greater treatment effects compared to those who were identified through a "one-step" process, either through one screening score or nomination. One example of a multiple gating process would be the use of Systematic Screening for Behavioral Disorders (SSBD) procedure. This process would necessitate the teacher to rank all students in the classroom and evaluate some of the highest ranking students with rating scales, followed by evaluation through structured observation by another staff (Walker & Severson, 2010). The identification process could affect treatment effects because accurately identifying students who really do need the intervention will allow them to truly reap the benefits of the intervention.

Currently, the most effective method for identifying internalizing problems in the schools is considered to be multiple gating procedures such as the Systematic Screening for Behavioral Disorders (SSBD) because they have the most empirical support (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007; Walker, Ramsey, & Gresham, 2004; Walker & Severson, 1990; Lane, Little, Casey, Lambert, Webby, Weisenbach & Phillips, 2009). On the other hand, there is evidence to suggest that using just office discipline referrals, teacher ratings, and teacher nomination are not very effective (Dowdy, Furlong, Eklund, Saeki, & Ritchey, 2010; Kamphaus, Thorpe, Winsor, Kroncke, Dowdy, VanDeventer, 2007; Kauffman, 1999; Lloyd, Kauffman, Landrum, &

Roe, 1991; Severson et al., 2007). Therefore, the finding from this moderator analysis is consistent with previous research.

Limitations

As with any systematic review, the resulting set of articles can vary dependent on the search terms used. Therefore, the use of keywords that are different from the ones used in this analysis could have initially generated a different set of potential articles. Consequently, the selected articles included in the analysis and the results of the analysis could have been different.

One limitation with identifying a mentor is that the definition is difficult to determine. Although the definition of a mentor and what constitutes a mentor in the context of this analysis has been discussed, it is important to note that there is still some variance. As discussed earlier, the definitions of a mentor are not identical in the literature (Dubois & Karcher, 2013; Hoyle et al., 2011; Rhodes, 2002). Thus, it may be an important step to more deeply examine the requirements and definitions of a mentor in the school setting as well as consider what would be the most beneficial for the students.

Another limitation to the study is that the vast majority of participants included in this synthesis were elementary school students in the general education setting without any identified disabilities. These characteristics limit the generalizability of the findings to students who are in middle and high schools as well as students who are identified with disabilities. Also, many studies did not identify characteristics of the mentors besides their role in the school setting. Therefore, a more in-depth analysis may necessitate the description of mentor variables.

For all participants included in this analysis, independent variables were manipulated and each phase (e.g., baseline and treatment phases) included a minimum of 3 data points. Also, except one study that used ABAB design (Christensen et al., 2007), the other seven studies used multiple baseline designs. These features in the studies are methodologically sound. However, the collection of interobserver agreement was not consistent across the studies. Five studies collected interobserver agreement for a least 20% of the data points with at least 80% agreement (Christensen et al., 2007; Dart et al., 2014; Marchant et al., 2007; Matthews, 2009; Ross et al., 2014) and one study for less than 16% of the data with 96% agreement (Hunter et al., 2014), while three studies did not report any interobserver data (Collin et al., 2016; Cook et al., 2015; Fiat et al., 2017).

The absence of criteria for interobserver reliability during the inclusion process is a limitation. Generally, broader methodological criteria (one that is more lenient), can lead to inclusion of studies that weaken the validity of the research synthesis (Lipsey & Wilson, 2001). Without reports of interobserver agreement, there is a possibility that behaviors may not be recorded with consistency due to problems with accuracy in judgment and incorrect recording of behaviors (Kazdin, 2011). Interobserver agreement helps ensure consistency between observers, minimize bias, and verify that goals were objectively defined (Kazdin, 2011). It is possible that having more strict criteria and limiting the number of studies leads to loss of data from excluded studies that could potentially be useful (Lipsey & Wilson, 2001). In these cases, it is important to examine the extent to which results differ based on the existence or absence of interobserver reliability data (Lipsey & Wilson, 2001). Sometimes, if the results of groups of studies

differ considerably, it could be beneficial to focus on the results of the studies that are considered better in quality. However, it is worth noting that the outcomes (i.e., effect sizes) for the studies with and without interobserver reliability were similar (see Table 1).

For analysis, Tau-U was used as it is considered a preferable method in single-case design studies (Parker et al., 2011; Vannest & Ninci, 2015), the results of single-case designs are regarded to be less generalizable compared to group designs. However, efforts were made to strengthen the internal and external validity of the study. First, one strength of single-case designs is that measurements are recorded over time, and the repeated measurements address internal validity problems with maturation and history. One way internal validity was strengthened was by establishing the stability of the baseline data so that the control is more effectively established. This was done by using Tau-U analysis and making baseline correction when the trend was significant. The external validity of the study was also strengthened by including only studies that establish a functional relationship and experimental control, providing a more convincing evidence of treatment effectiveness. Also, efforts were made to present all information regarding the studies that were able to be extracted regarding student, mentor, and intervention characteristics.

Additionally, although the synthesis examined interventions that addressed internalizing problems, the specific outcome measures did not reflect one type of internalizing behavior and some did not seem to accurately reflect an internalizing problem. There may be some problems in translating the identified problems to measurable goals. For instance, monitoring social withdrawal by measuring percent

participation may not be ideal. If there are such differences, it would be important for future studies to justify the outcomes that were selected to be monitored.

With the moderator analysis, it is important to note that they were evaluated by simply examining the differences in the Tau-U mean effect sizes without the use of statistical comparisons like significance tests. Furthermore, because the confidence intervals overlap between the levels of each moderator, it is difficult to determine with certainty that there was a difference between the levels. This may contribute to why, in the moderator analysis, interventions with higher fidelity data did not produce greater effects as expected. Future studies should take the statistical limitations into consideration and conduct meta-analyses that address them by utilizing d-statistic or g-statistic as described by Shadish, Hedges, and Pustejovsky (2014).

Currently, empirical support for mentor-based interventions in addressing internalizing problems is not well-established. However, the analysis does provide some insight to the intervention's overall effectiveness, possible moderating factors of the intervention, and directions for future studies.

Conclusions and Implications

The results of the analyses could be helpful to educators who may consider mentor-based intervention for addressing internalizing problems. This synthesis provides some insight that these interventions could be used within the schools.

The strength of this intervention seems to be that there is flexibility in who can serve as mentors. In their study, Hunter and colleagues (2014) indicated that teachers were able to implement the intervention without training in cognitive-behavioral

interventions even though the intervention did incorporate some components of CBT. Although this does not necessarily indicate that there should be no training involved, it suggests that perhaps there is no need for extensive training. Cook and colleagues (2012) also noted that the intervention was not designed to be delivered by mentors who have specialized training or mental health credential. It is not just these two studies that support this idea. Throughout the studies, the mentors varied widely (i.e., student teacher, school psychology graduate student, school psychologist, school counselor, special education teacher, paraprofessional, undergraduate majoring in psychology, and undergraduate majoring in special education). Despite this variation, the overall outcome suggest that these interventions yield positive effects. Therefore, further studies may help to confirm that mentors do not need to have an advanced knowledge of mental health to address less severe internalizing problems.

Also, it may be appropriate to consider mentor-type interventions for addressing mild to moderate internalizing problems. Specific recommendations have been made stating that these interventions are appropriate for less severe problems (Cook et al., 2015; Fiat et al., 2017; Herrera & Karcher, 2011). One reason for this recommendation likely has to do with the preparation of mentors and their ability to handle certain levels of challenges with the students. Accordingly, it should be the noted that many counseling-type interventions are used to treat anxiety and depression, conducted one-to-one a one to one basis, and are generally considered an intensive service or a Tier 3 service within the MTSS framework. On the other hand, the studies included in this research synthesis consider the interventions as prevention-based interventions or Tier 2

interventions. Therefore, none of the students included in the studies were diagnosed with depression, anxiety, or any psychiatric disorders. Rather, students were either at-risk for internalizing disorders or were identified with internalizing problems without diagnoses. Perhaps these interventions can be used as stand-alone interventions for mild to moderate problems and additional treatment for more severe ones.

With the existence of such empirically supported intervention like CBT, it may not be appropriate to consider mentor-type interventions as a replacement of counseling-type interventions. Mentor-type interventions may be considered as an alternative for less severe or additional treatment for more severe internalizing problems. This study, to some extent, provides insight on the benefits of these interventions.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Atkins, M. S., Frazier, S. L., Adil, J. A., & Talbott, E. (2003). School-based mental health services in urban communities. In M. Weist, S. Evans, & N. Tashman (Eds.), *Handbook of school mental health: Advancing practice and research* (pp. 165−178). New York: Kluwer Academic/Plenum Publishers.

Averill, O. H., & Rinaldi, C. (2011). Multi-tier system of supports. *District Administration*, *47*(8), 91.

Barrera, M., & D., Bonds (2013). Mentoring relationships and social support. In D. L., Dubois, & M. J., Karcher (Eds.), *Handbook of youth mentoring.* Sage Publications.

Barrett, P. M., Farrell, L., Pina, A. A., Peris, T. S., & Piacentini, J. (2008). Evidence-based psychosocial treatments for child and adolescent obsessive–compulsive disorder. *Journal of Clinical Child & Adolescent Psychology*, *37*(1), 131-155.

Beck, A. T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy*, *1*, 184-200.

Beesdo, K., Bittner, A., Pine, D. S., Stein, M. B., Lieb, R., & Wittchen, H. U. (2007). Incidence of social anxiety disorder and the consistent risk for secondary depression in the first three decades of life. *Archives of General Psychiatry*, 64, 903–912

Bennett, D. S., & Gibbons, T. A. (2000). Efficacy of child cognitive-behavioral interventions for antisocial behavior: A meta-analysis. *Child & Family Behavior Therapy*, *22*(1), 1-15.

Bernstein, L., Rappaport, C. D., Olsho, L., Hunt, D., & Levin, M. (2009). Impact Evaluation of the US Department of Education's Student Mentoring Program. Final Report. NCEE 2009-4047. *National Center for Education Evaluation and Regional Assistance*.

Bradshaw, C. P., Buckley, J. A., & Ialongo, N. S. (2008). School- based service utilization among urban children with early onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly,* 23, 169–186.

Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., & Erkanli, A. (1995). Children's mental health service use across service sectors. *Health affairs*, *14*(3), 147-159.

Christensen, L., Young, K. R., & Marchant, M. (2007). Behavioral intervention planning: Increasing appropriate behavior of a socially withdrawn student. *Education and Treatment of Children*, *30*(4), 81-103.

Cicchetti, D., & Toth, S. L. (Eds.) (1991) *Internalizing and Externalizing Expressions of Dysfunction*. Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hilsdale. *NJ: Lawrence Earlbaum Associates*, *2*.

Collins, T. A., Gresham, F. M., & Dart, E. H. (2016). The effects of Peer-Mediated Check-In/Check-Out on the social skills of socially neglected students. *Behavior modification*, *40*(4), 568-588.

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

Cook, C. R., Xie, S. R., Earl, R. K., Lyon, A. R., Dart, E., & Zhang, Y. (2015). Evaluation of the courage and confidence mentor program as a Tier 2 intervention for middle school students with identified internalizing problems. *School Mental Health*, *7*(2), 132-146.

Costello, E. J., Mustillo, S., Erkanil, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry,* 60(8), 837–844.

Crone, D. A., Hawken, L. S., & Horner, R. H. (2010). *Responding to problem behavior in schools: The behavior education program*. Guilford Press.

Dart, E. H., Collins, T. A., Klingbeil, D. A., & McKinley, L. E. (2014). Peer management interventions: A meta-analytic review of single-case research. *School Psychology Review*, *43*, 367–384.

Dart, E. H., Furlow, C. M., Collins, T. A., Brewer, E., Gresham, F. M., & Chenier, K. H. (2015). Peer-mediated check-in/check-out for students at-risk for internalizing disorders. *School Psychology Quarterly*, *30*(2), 229.

David-Ferdon, C., & Kaslow, N. J. (2008). Evidence-based psychosocial treatments for child and adolescent depression. *Journal of Clinical Child & Adolescent Psychology*, *37*(1), 62-104.

Dowdy, E., Furlong, M., Eklund, K., Saeki, E., & Ritchey, K. (2010). Screening for mental health and wellness: Current school based practice and emerging possibilities. In B. Doll, W. Pfohl, & J. Yoon (Eds.), *Handbook of youth prevention science*. New York: Routledge.

DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American journal of community psychology*, *30*(2), 157-197.

DuBois, D. L., & Karcher, M. J. (Eds.). (2013). *Handbook of youth mentoring*. Sage Publications.

DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest*, *12*(2), 57-91.

Eisenberg, N., Cumberland, A., Spinrad, T. L., Fabes, R. A., Shepard, S. A., Reiser, M., ... & Guthrie, I. K. (2001). The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Development*, *72*(4), 1112-1134.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Farmer, E. M., Burns, B. J., Phillips, S. D., Angold, A., & Costello, E. J. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric services*, *54*(1), 60-66.

Fiat, A. E., Cook, C. R., Zhang, Y., Renshaw, T. L., DeCano, P., & Merrick, J. S. (2017). Mentoring to promote courage and confidence among elementary school students with internalizing problems: A single-case design pilot study. *Journal of Applied School Psychology*, 1-27.

Forns, M., Abad, J., & Kirchner, T. (2011). Internalizing and externalizing problems. In *Encyclopedia of adolescence* (pp. 1464-1469). New York, NY: Springer.

Ginsburg, G. S., & Drake, K. L. (2002). School-based treatment for anxious African-American adolescents: A controlled pilot study. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(7), 768–775.

Greenberg, M., Domitrovich, C., & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention & Treatment*, *4*(1), 1–59.

Gresham, F. M., & Kern, L. (2004). Internalizing behavior problems in children and adolescents. In R., Rutherford, M. Quinn, & S. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 262–281). New York: Guilford Press.

Grills, A. E., & Ollendick, T. H. (2002). Peer victimization, global self-worth, and anxiety in middle school children. Journal of Clinical and Adolescent Psychology, 31(1), 59–68.

Hawken, L. S. (2006). School psychologists as leaders in the implementation of a targeted intervention: The Behavior Education Program. *School Psychology Quarterly*, *21*(1), 91.

Hawken, L. S., Bundock, K., Kladis, K., O'Keeffe, B., & Barrett, C. A. (2014). Systematic review of the check-in, check-out intervention for students at risk for emotional and behavioral disorders. *Education and Treatment of Children*, *37*(4), 635-658.

Hawken, L. S., & Horner, R. H. (2003). Evaluation of a targeted intervention within a schoolwide system of behavior support. *Journal of Behavioral Education*, *12*(3), 225-240.

Hawken, L. S., MacLeod, K. S., & Rawlings, L. (2007). Effects of the behavior education program (BEP) on office discipline referrals of elementary school students. *Journal of Positive Behavior Interventions*, *9*(2), 94-101.

Herrera, C., & Karcher, M. (2013). School-based mentoring. In D. L., Dubois, & M. J., Karcher (Eds.), *Handbook of youth mentoring.* Sage Publications.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111*, 127–155.

Hops, H., Finch, M., & McConnell, S. (1985). Social skills deficits. In P. Bornstein & A. Kazdin (Eds.), *Handbook of clinical behavior therapy with children* (pp. 543-598). Homewood, IL: The Dorsey Press.

Horn, W., & Packard, T. (1985). Early identification of learning problems: A meta-analysis. *Journal of Educational Psychology,* 77(5), 597-607.

Hoyle, C. G., Marshall, K. J., & Yell, M. L. (2011). Positive behavior supports: Tier 2 interventions in middle schools. *Preventing School Failure, 55*(3), 164-170.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Humensky, J., Kuwabara, S. A., Fogel, J., Wells, C., Goodwin, B., & Van Voorhees, B. W. (2010). Adolescents with depressive symptoms and their challenges with learning in school. *The Journal of School Nursing, 26*, 377–392.

Hunter, K. K., Chenier, J. S., & Gresham, F. M. (2014). Evaluation of Check In/Check Out for students with internalizing behavior problems. *Journal of Emotional and Behavioral Disorders*, *22*(3), 135-148.

Ingram, R., & Smith, L. T. (2008). Mood disorders. In J. E. Maddux & B. A. Winstead (Eds.), *Psychopathology: Foundations for a contemporary understanding* (2nd ed., pp. 191–197). New York, NY: Routledge.

Ishikawa, S. I., Okajima, I., Matsuoka, H., & Sakano, Y. (2007). Cognitive behavioural therapy for anxiety disorders in children and adolescents: A meta-analysis. *Child and Adolescent Mental Health*, *12*(4), 164-172.

Kamphaus, R. W., Thorpe, J. S., Winsor, A. P., Kroncke, A. P., Dowdy, E. T., & VanDeventer, M. C. (2007). Development and predictive validity of a teacher screener for child behavioral and emotional problems at school. *Educational and Psychological Measurement, 67*, 342–356. doi: 10.11770013164406292041

Kauffman, J. M. (1999). How we prevent the prevention of emotional and behavioral disorders. Exceptional Children, 65(4), 448−468.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. Oxford University Press.

Kelder, S. H., Murray, N. G., Orpinas, P., Prokhorov, A., McReynolds, L., Zhang, Q., & Roberts, R. (2001). Depression and substance use in minority middle-school students. *American Journal of Public Health, 91*, 761–766.

Kendall, P. C., & Hedtke, K. A. (2006). *Cognitive-behavioral therapy for anxious children: Therapist manual*. Workbook Publishing.

Kerr, M., Tremblay, R. E., Pagani, L., & Vitaro, F. (1997). Boys' behavioral inhibition and the risk of later delinquency. *Archives of General Psychiatry,54*(9), 809-816.

Kessler, R. C., Berglund, P. A., Demler, O., Jin, R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry*, 62, 593–602.

Kim, J. S. (2008). Examining the effectiveness of solution-focused brief therapy: A meta-analysis. *Research on Social Work Practice*, *18*(2), 107-116.

Kovacs, M., & Devlin, B. (1998). Internalizing disorders in childhood. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *39*, 47–63.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. *What Works Clearinghouse*.

Lane, K. L., Little, M. A., Casey, A. M., Lambert, W., Wehby, J., Weisenbach, J. L., & Philips, A. (2009). A comparison of systematic screening tools for emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 17*, 93–105. doi:10.1177/1063426608326203

Last, C. G., Perrin, S., Hersen, M., & Kazdin, A. E. (1996). A prospective study of childhood anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(11), 1502-1510.

Levitt, V. H., & Merrell, K. W. (2009). Linking assessment to intervention for internalizing problems of children and adolescents. *School Psychology Forum: Research in Practice*, 3, 13–26.

Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological per spectives on the integration of affect, motivation, and cognition. *Educational Psychology Review*, *18*, 307–314.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: Sage publications.

Lloyd, J. W., Kauffman, J. M., Landrum, T. J., & Roe, D. L. (1991). Why do teachers refer pupils for special education? An analysis of referral records. Exceptionality, 2(3), 115−126.

Maag, J. W., & Swearer, S. M. (2005). Cognitive-behavioral interventions for depression: Review and implications for school personell. *Behavioral Disorders*, *30*(3), 259-276.

Maggin, D. M., Swaminathan, H., Rogers, H. J., O'Keeffe, B. V., Sugai, G., & Horner, R. H. (2011). A generalized least squares regression approach for computing effect sizes in single-case research: Application examples. *Journal of School Psychology, 49*(3), 301-321.

Maggin, D. M., Zurheide, J., Pickett, K. C., & Baillie, S. J. (2015). A systematic evidence review of the check-in/check-out program for reducing student challenging behaviors. *Journal of Positive Behavior Interventions, 17*(4), 197-208.

Mahoney, M. J. (1974). *Cognition and behavior modification*. Cambridge, MA: Ballinger.

March, R. E., & Horner, R. H. (2002). Feasibility and contributions of functional behavioral assessment in schools. *Journal of Emotional and Behavioral Disorders, 10*(3), 158-170.

Marchant, M. R., Solano, B. R., Fisher, A. K., Caldarella, P., Young, K. R., & Renshaw, T. L. (2007). Modifying socially withdrawn behavior: A playground intervention for students with internalizing behaviors. *Psychology in the Schools, 44*(8), 779.

Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradović, J., Riley, J. R., Kristen, B. & Tellegen, A. (2005). Developmental cascades: linking academic achievement and externalizing and internalizing symptoms over 20 years. *Developmental psychology, 41*(5), 733.

McCary, C. A. & Weisz, J.r. (2007). Effects of psychotherapy for depression in children and adolescents: What we can (and can't) learn from meta analysis and component profiling. *Journal of American Academy of Child Adolescent Psychiatry,* 46 (7), 879-886.

Meichenbaum, D. (1977). *Cognitive-behavior modification.* New York: Plenum.

Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the national comorbidity survey-adolescent suppl. (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 49,* 980 –989.

Merrell, K. W., & Gueldner, B. A. (2010). Preventive interventions for students with internalizing disorders: Effective strategies for promoting mental health in schools. In M. R. Shinn & H. M. Walker (Eds.), *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 799–824). Bethesda, MD: National Association of School Psychologists.

Meyers, A. W., & Craighead, W. E. (1984). Cognitive behavior therapy with children. In *Cognitive behavior therapy with children* (pp. 1-17). Springer US.

Miller, D. N. and Nickerson, A. B. (2007) 'Projective Techniques and the School-Based Assessment of Childhood Internalizing Disorders: A Critical Analysis', *Journal of Projective Psychology and Mental Health* 14: 48–58.

Muroff, J., & Ross, A. (2011). Social disability and impairment in childhood anxiety. In D. McKay & E. A. Storch (Eds.), *Handbook of child and adolescent anxiety disorders* (pp. 457–478). New York, NY: Springer.

Myers, D. M., & Briere III, D. E. (2010). Lessons Learned from Implementing a Check-in/Check-out Behavioral Program in an Urban Middle School.*Beyond Behavior*, *19*(2), 21-27.

Parker, R. I., & Hagan-Burke, S. (2007). Useful effect size interpretations for single case research. *Behavior Therapy*, *38*(1), 95-105.

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, *35*(4), 303-322.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284-299.

Plotts, C. A., & Lasser, J. (2013). *School psychologist as counselor: A practitioner's handbook*. National Association of School Psychologists.

Portwood, S. G., Ayers, P. M., Kinnison, K. E., Waris, R. G., & Wise, D. L. (2005). YouthFriends: Outcomes from a school-based mentoring program. *The Journal of Primary Prevention*, *26*(2), 129-188.

Reynolds, W. M. (1992). *The study of internalizing disorders in children and adolescents*. John Wiley & Sons.

Richards, H. J., & Hadwin, J. A. (2011). An exploration of the relationship between trait anxiety and school attendance in young people. *School Mental Health*, *3*, 236–244.

Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review, 3,* 223–241.

Rones, M., & Hoagwood, K. (2001). School-based mental health services: A 15-year research review. *Child Clinical and Family Psychology Review*, 3, 223−241.

Ross, S. W., & Sabey, C. V. (2015). Check-in check-out+ social skills: Enhancing the effects of check-in check-out for students with social skill deficits. *Remedial and Special Education*, *36*(4), 246-257.

Schwartz, S. O., Rhodes, J. E., & Herrera, C. (2012). The influence of meeting time on academic outcomes in school-based mentoring. *Children and Youth Services Review*, *34*, 2319–2326.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-sub- ject research: Issues and applications. *Behavior Modi cation, 22*, 221–242.

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

Shadish, W. R., Hedges, L. V., & Pustejovsky, J. E. (2014). Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *Journal of School Psychology*, *52*(2), 123-147.

Silverman, W. K., Pina, A. A., & Viswesvaran, C. (2008). Evidence-based psychosocial treatments for phobic and anxiety disorders in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *37*(1), 105-130.

Skinner. B. F. (1963). Operant behavior. *American Psychologist, 18*, 503-515.

Spielmans, G. I., Pasek, L. F., & McFall, J. P. (2007). What are the active ingredients in cognitive and behavioral psychotherapy for anxious and depressed children? A meta-analytic review. *Clinical psychology review*, *27*(5), 642-654.

Tarlow, K. R. (2016). An Improved Rank Correlation Effect Size Statistic for Single-Case Designs Baseline Corrected Tau. *Behavior modification*.

Todd, A., Benedict, E. A., Horner, R. H., McKenna, M. K., Filter, K. J., & Watson, J. (2007). Check in/check out: A post-hoc evaluation of an efficient, secondary-level targeted intervention for reducing problem behaviors in schools. *Education and Treatment of Children*, *30*(1), 69-84.

Tolan, P., Henry, D., Schoeny, M., Bass, A., Lovegrove, P., & Nichols, E. (2013). Mentoring interventions to affect juvenile delinquency and associated problems: A systematic review. *Campbell Systematic Reviews*, *9*(10).

Vannest, K. J., & Ninci, J. (2015). Evaluating Intervention Effects in Single-Case Research Designs. *Journal of Counseling & Development*, *93*(4), 403-411.

Walker, H. M., Nishioka, V. M., Zeller, R., Severson, H. H., & Feil, E. G. (2000). Causal factors and potential solutions for the persistent under- identification of students having emotional or behavioral disorders in the context of schooling. *Assessment for Effective Intervention, 26*, 29–39.

Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices*, 2nd ed. Belmont, CA: Wadsworth/Thomson Learning.

Walker, H. M., & Severson, H. H. (1990). *Systematic Screening for Behavior Disorders (SSBD)*. Longmont, CO: Sopris West.

Walker, H. M., & Severson, H. H. (1994). Replication of the Systematic Screening for Behavior Disorders (SSBD) procedure for the identification of at-risk children. Journal of Emotional and Behavioral Disorders, 2(2), 1063−4266.

Walker, H. M., Severson, H., Stiller, B., Williams, G., Haring, N., Shinn, M., & Todis, B. (1988). Systematic screening of pupils in the elementary age range at risk for behavior disorders development and trial testing of a multiple gating model. *Remedial and Special Education, 9*(3), 8-20.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially," all must have prizes.".

Watson, J. (1930). *Behaviorism*. New York: W. W. Norton.

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychological bulletin, 132*(1), 132.

Wilson, G. T. (1982). Psychotherapy process and procedure: The behavioral mandate. *Behavior Therapy*, *13*(3), 291-312.

Winett, R. A., & Winkler, R. C. (1972). Current behavior modification in the classroom: Be still, be quiet, be docile. *Journal of Applied Behavior Analysis*, *5*(4), 499-504.

Wolery, M., Busick, M., Reichow, B., & Barton, E. E. (2010). Comparison of overlap methods for quantitatively synthesizing single-student data. *The Journal of Special Education, 44*(1), 18-28.

Wolfe, K., Pyle, D., Charlton, C. T., Sabey, C. V., Lund, E. M., & Ross, S. W. (2016). A systematic review of the empirical support for Check-In Check-Out. *Journal of Positive Behavior Interventions, 18*(2), 74-88.

World Health Organization. (2004). Prevention of mental disorders: Effective interventions and policy options. Summary Report. Geneva: Author.

World Health Organization. (2012). International statistical classification of diseases and related health problems. Geneva: World Health Organization.

Zahn–Waxler, C., Klimes–Dougan, B., & Slattery, M. J. (2000). Internalizing problems of childhood and adolescence: Prospects, pitfalls, and progress in understanding the development of anxiety and depression. *Development and psychopathology*, *12*(3), 443-466.

Appendix

Table 1
Single Subject Design Studies

Authors	Year	N	DV	IV	Tau-U	P-Value	CI 90%
Hunter et al.	2014	4	Percentage on DPR	CICO (adjusted)	**0.881**	0.000	0.624-1
Cook et al.	2015	4	Distress rating on SUD	CCMP	**0.795**	0.000	0.554-1
Fiat et al.	2017	6	Percentage on DBR	CCMP	**0.817**	0.000	0.595-1
Marchant et al.	2007	3	Rate of effective communication	Social Skill	**0.993**	0.000	0.778-1
Dart et al.	2015	3	Points on DBR	CICO	**0.658**	0.000	0.379-1
Christensen et al.	2007	1	Percentage of socially appropriate behavior	Social Skill	**0.792**	0.000	0.460-1
Ross & Sabey	2015	4	Percentage of positive social engagement	CICO (adjusted)	**0.853**	0.000	0.552-1
Collins et al.	2016	4	Points on DBRC	CICO	**0.729**	0.000	0.454-1

Table 2
Main Effects for Single Subject Design Studies as a Function of Sample Population, Mentor Characteristic, and Intervention

Moderators	Levels	N	Tau-U	P-Value	CI 90%
Student Grade	Preschool-3rd	7	**0.776**	0.000	0.557-0.995
	4th-6th	18	**0.847**	0.000	0.697-0.998
	7th-High School	4	**0.795**	0.000	0.508-1.000
Student Disability	Identified	4	**0.735**	0.000	0.473-0.998
	Not Identified	24	**0.835**	0.000	0.732-0.938
Student Race/Ethnicity	Caucasian	13	**0.875**	0.000	0.730-1.000
	Minority	16	**0.778**	0.000	0.649-0.906
Method for Student Selection	Multiple Gating	14	**0.857**	0.000	0.726-0.987
	Screening/Nomination	15	**0.785**	0.000	0.648-0.925
Mentor Type		21	**0.864**	0.000	0.751-0.977
	Adult	8	**0.709**	0.000	0.527-0.892
	Peer				
Mentor's Mental Health Background		4	**0.881**	0.000	0.634-1.000
	Yes	12	**0.752**	0.000	0.595-0.910
	No				
Meeting Frequency		22	**0.824**	0.000	0.715-0.934
	3-5 per Week	7	**0.813**	0.000	0.611-1.000
	1-2 per Week				
Intervention Length		21	**0.808**	0.000	0.698-0.917
	More than 10 Days	8	**0.862**	0.000	0.665-1.000
	10 Days or Less				
Treatment Fidelity		22	**0.785**	0.000	0.674-0.895
	> 80%	10	**0.813**	0.000	0.666-0.960
	< 80% or Not Collected				
CICO Variation		7	**0.697**	0.000	0.499-0.894
	Without Modification	18	**0.833**	0.000	0.706-0.960
	With Modification to Internalizing Problem				

Table 3
Study Details

Study	Mentor Title	Student Selection	Independent Variable (Intervention information)
Hunter et al. (2014)	School psychology graduate student	Teacher nomination	CICO adjusted: During "check-in," problems experienced on the previous day was discussed and solved. When needed, the two worked to identify and replace maladaptive thoughts with productive ones. During "check-out", rewards were awarded if student met daily goal. If not met, mentor helped student problem-solve to perform better the next day.
Cook et al. (2015)	School psychologist and special education resource teacher	Screening, then ratings from student	Courage and Confidence Mentor Program (CCMP): Initially, two 40-minute sessions were conducted. First session was designed to build rapport with student, raise emotional awareness, provide students with psychoeducation to normalize and externalize negative emotions, and teach coping skills. The second session was designed to review previous content, teach "courage tools" such as deep breathing and using guided imagery, and teach how to use courageous self-statements. Check in process was used to encourage the student and pre-correct problems, while check out process was used to facilitate positive interaction and provide performance feedback.
Fiat et al. (2017)	School counselor and paraprofessional	Screening, then ratings from student	Courage and Confidence Mentor Program (CCMP): Same as CCMP mentioned above.

Table 4
Study Details Continued

Study	Mentor Title	Student Selection	Independent Variable (Intervention information)
Marchant et al. (2007)	Undergraduate majoring in psychology, undergraduate majoring in special education, and school psychology graduate student	Teacher checklist completion, school service team nomination, then informal observation	Social skills training: Steps of social skill were taught, modeled, and practiced. Mentor encouraged student to use social skills with other students, provided daily reinforcement using point system, and checked in with student daily.
Dart et al. (2015)	Peer	Screening	Peer-mediated CICO: Standard CICO with peer as mentor.
Christensen et al. (2007)	Peer	Direct observation, then SSBD	Social skills: Mentor observed and evaluated student's classroom, provided feedback and reinforcement, and conversed with the student to increase social interaction.
Ross & Sabey (2015)	Student teacher	Teacher and principal nomination, then SSRS by teacher	CICO with social skills component: Implementation of the social skills component took 15 minutes each day. This time was used for explicit social skills instruction, modeling, and guided practice. Student was also assigned homework to further practice new skills. This component of the intervention lasted anywhere from a few days to two weeks depending on adequacy of student performance.
Collins et al. (2016)	Peer	Screening, then nomination	Peer-mediated CICO: Standard CICO with peer as mentor.

Table 5
Participant Details

Authors	N	Name	Gender	Grade	Ethn-icity	Dependent Variable (Measured behavior)
Hunter et al. (2014)	1	Patrick	Male	4th	AA	"Pays attention to the lesson and focuses on completing classwork."
	2	Chris	Male	4th	C	"Makes eye contact with others when speaking to them."
	3	Caroline	Female	4th	C	"Participates appropriately in the class activity."
	4	Jeff	Male	4th	C	"Is on task, requiring no more than three redirections." (p. 140-141)
Cook et al. (2015)	5	Ashley	Female	7th	C	Subjective Units of Distress (SUD) rating assessing internalizing symptoms.
	6	John	Male	6th	A	
	7	Megan	Female	8th	C	
	8	Ben	Male	8th	A	
Fiat et al. (2017)	9	Arturo	Male	4th	H	Percent participation
	10	Jamilla	Female	5th	AA	Percent participation
	11	Sue	Female	4th	A	Percent participation
	12	Marcus	Male	4th	AA	Frequency of somatic complaints
	13	Jamie	Female	5th	C	Percent participation
	14	Juan	Male	4th	H	Percent participation
Marchant et al. (2007)	15	Catherine	Female	1st	C	Rate of effective communication—defined as appropriately engaging a peer by looking at him or her and initiating verbal communication, physical gestures such as waving and giving a thumbs up, high five, or handshake
	16	Michael	Male	5th	C	
	17	Scott	Male	5th	C	

Note. AA=African American, C=Caucasian, A=Asian, H=Hispanic. For dependent variables that were measured by DPR, DBR, or DBRC, just one of the target behaviors is provided in the table as an example.

Table 6
Participant Details Continued

Authors	N	Name	Gender	Grade	Ethn-icity	Dependent Variable (Measured behaviors)
Dart et al. (2015)	18	Stephanie	Female	1st	C	"Spent time with other students."
	19	Sarah	Female	2nd	AA	"Appeared outgoing and sociable."
	20	Christina	Female	1st	AA	"Smiled and appeared happy." (p. 234)
Christensen et al. (2007)	21	Jose	Male	3rd	H	Percentage of socially appropriate behavior. Some examples include attending to the teacher, answering questions when requested, and complying with teacher instruction.
Ross & Sabey (2015)	22	Lucinda	Female	3rd	H	Percentage of positive social engagement defined as appropriate play or positive communication with peers ranging from neutral to complimentary. These behaviors include statements of approval, negotiations, and appropriate complaints.
	23	Sarah	Female	3rd	W	
	24	Emily	Female	5th	H	
	25	Olivia	Female	5th	W	
Collins et al. (2016)	26	Madeleine	Female	5th	C	"Starting conversations with peers."
	27	Ferdinand	Male	5th	AA	Ignoring classmates who attempt to distract student.
	28	Taylor	Female	5th	AA	"Interacting appropriately with peers".
	29	Gertrude	Female	4th	AA	"Asking for help when needed." (p. 573-574)

Note. AA=African American, C=Caucasian, A=Asian, H=Hispanic. For dependent variables that were measured by DPR, DBR, or DBRC, just one of the target behaviors is provided in the table as an example.