Protocol for a Systematic Review: Self-Management Interventions for Reducing Challenging Behaviors among School-age Students: A Systematic Review

Aaron M. Thompson, Lindsay R. Ruhr, Brandy R. Maynard, Michael D. Pelts, Natasha K. Bowen

Submitted to the Coordinating Group of:

☐	Crime and Justice
☒	Education
	☐ Disability
☐	International Development
	☐ Nutrition
☐	Social Welfare
☐	Other:

Plans to co-register:

☒	No		
☐	Yes	☐ Cochrane	☐ Other
☐	Maybe		

Date Submitted: 25 March 2013
Date 1st Revision Submitted: 18 September 2013
Date 2nd Revision Submitted: 2 October 2013
Approval Date:
Publication Date: 1 November 2013

BACKGROUND

The Problem, Condition or Issue

Most parents and teachers agree that students need to exhibit appropriate social behaviors in order to achieve academic goals; however, approximately 20% of students, or 3-4 students in the average classroom, repeatedly display challenging behaviors that interfere with normal academic and social development (Brauner & Stephens, 2006; Bushaw & Lopez, 2010; Satcher, 2004; Walker, Ramsey, & Gresham, 2004). Challenging student behaviors are defined as acts that (a) interfere with social and academic functioning; (b) harm a child, his or her peers, or adults; and (c) place a child at risk for developmental problems. Because research often distinguishes between subtypes of challenging behaviors, we specify three broad subtypes: (a) direct and indirect forms of aggression (e.g., hitting, name calling, spreading rumours; Dodge & Coie, 1987; Leff & Crick, 2010; Parke & Slaby, 1983); (b) overt and covert antisocial behaviors (e.g., stealing, bullying, lying, cheating); and (c) low intensity acts of insubordination (e.g., noncompliance, withdrawal, refusal to cooperate, impulsivity, inattention, off-task; Kaiser & Rasminsky, 2009).

Challenging student behaviors are harmful to everyone in schools—including students who exhibit the behaviors and their peers and teachers. Students who exhibit challenging behaviors are frequently removed from class, which interrupts instruction, exacerbates academic difficulties, and increases the likelihood of school failure and dropout (Gresham, Lane, & Lambros, 2000; Nelson, Benner, Lane, & Smith, 2004). Peers of disruptive students are adversely affected by the behaviors due to lost instructional opportunities (U.S. Department of Education [USDOE], 2006). Observational studies indicate challenging behaviors contribute to a loss of four hours of instruction per week in the average classroom or about 144 hours per student over the academic year (Walker et al., 2004). Lastly, teachers experience increased stress and burnout associated with managing challenging behaviors (Brouwers & Tomic, 2000; Clunies- Ross, Little, & Kienhuis, 2008; Grayson & Alvarez, 2008; Hastings & Bham, 2003; Joseph & Strain, 2003). A survey of highly-qualified teachers suggested that 53% of those who requested transfers and 44% of those who quit teaching cited challenging student behaviors as their primary reason for the decision (USDOE, 2005). Because challenging behaviors adversely impact everyone in schools, it is vital that school professionals assist students with challenging behaviors to learn adaptive social, emotional, and behavioral skills.

Research suggests school-based programs that promote competencies in social, emotional, and behavioral skills hinge on the development of five interrelated concepts: social-awareness, self-awareness, self-management, relationship, and problem solving (Bridgeland, Bruce, & Hariharan, 2013). Also called social, emotional, noncognitive, or soft skills, exposure to these skills increases the likelihood that students with challenging behaviors will experience better proximal school-related and distal life-course outcomes (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Heckman & Kautz, 2012; Wilson & Lipsey,

2007). For example, programs and practices promoting the development of the aforesaid skills are related to improvements in social functioning (ES = 0.69), attitudes toward school (ES = 0.24), behavioral problems (ES = 0.26), emotional stability (ES = 0.28), and academic performance (ES = 0.28; Durlak et al., 2011).

To facilitate the development of social, emotional, and behavioral skills, researchers and educators increasingly recognize the importance of autonomy support as an intervention mechanism (Algozzine, Browder, Karvonen, Test, & Wood, 2001; Field, Martin, Miller, Ward, & Wehmeyer, 1998; Lane, Menzies, Bruhn, & Crnobori, 2010). Autonomy refers to a sense of self-management (Deci & Cascio, 1972; Deci, Cascio, & Krusell, 1975; Deci & Ryan, 2011; Wigfield, Eccles, Roeser, & Schiefele, 2008; Wigfield, Eccles, Schiefele, Roeser, & Davis- Kean, 2007). Authority figures who endorse and enable the development of skills and opportunities required for self-management engage in autonomy support strategies (Deci & Ryan, 2011; Field et al., 1998). Strategies that integrate principles of autonomy support include—but are not be limited to—instruction in decision making, problem solving, goal setting, self-awareness, self-assessment, self-evaluation, self-management, and self-monitoring (Algozzine et al., 2001; Lane et al., 2010; Wehmeyer & Schwartz, 1997). Autonomy support strategies also facilitate improved student-teacher relations (Wentzel, Filisetti, & Looney, 2007). Improved student-teacher relations diminishes challenging behaviors and makes disciplining students who display those behaviors more effective (Hamre & Pianta, 2003).

In summary, promoting the development of competencies in social, emotional, and behavioral skills requires on-going and quality instruction in self-awareness, social awareness, self-management, relationship, and problem solving skills. Educators can nurture and cultivate the development of these valuable skills through autonomy support strategies that propagate an increased sense of self-management in students. Though many strategies impart social emotional skills and promote student autonomy separately, a self-management intervention combines social and emotional skills with autonomy support in a single approach.

Self-Management Interventions

The review will focus on the effectiveness of school-based *self-management* (SM) interventions—a widely-used cognitive behavioral intervention to address disruptive and challenging behaviors in school settings. Though SM interventions are referred by many names (e.g., self-control, effortful control, self-regulation), SM is defined as a set of strategies that students are trained in to assess, monitor, and evaluate their own behavioral performance (Briesch & Chafouleas, 2009; Cole, Marder, & McCann, 2000; Fantuzzo, Polite, Cook, & Quinn, 1988; Maggin, Briesch, & Chafouleas, 2013; Rothbart & Rueda, 2005; Shapiro & Cole, 1994; Shapiro, Durnan, Post, & Skibitsky-Levinson, 2002). More specifically, Fantuzzo and colleagues (1988) suggest a SM intervention includes one or a combination the following elements:

1. self-selecting a target behavior
2. self-defining the target behavior
3. self-determining a performance goal
4. self-identifying reinforcers
5. self-prompting a reflection of behavior
6. self-observing a target behavior
7. self-recording the observations
8. self-charting the observations
9. self-appraising performance
10. self-administering primary reinforcers
11. self-administering secondary reinforcers

One challenge faced in the proposed review is that researchers and practitioners rarely implement the elements of SM in a consistent manner. The SM literature has been summarized by five prior reviews. The first review applied the above classification to all SM studies conducted between 1967 and 1988 and noted that studies included 9.6 out of the 11 elements listed above (88%)—with 60% of studies including all 11 elements (Fantuzzo et al., 1988). In a second follow-up review of all SM studies published between 1989 and 2008, Briesch and Chafouleas (2009) applied the same typology used by Fantuzzo and found that 7.6 out of the 11 elements listed above (60%)—with only 3% of the studies including all 11 elements. Upon closer examination, Briesch and Chafouleas noted 4 out of the 11 SM elements "consistently" appeared across the 30 SM studies (i.e., self-selecting a target behavior, self-defining the target behavior, self-observing the target behavior, and self-recording the observations) and a fifth step (i.e., self-evaluation) was noted in half of the studies reviewed (Briesch & Chafouleas, 2009, p. 111).

A third review examined the impact of SM on academic outcomes and proposed that a SM intervention consisted of five elements: i.e., goal setting, self-monitoring, self-evaluation, self-instruction, and strategy instruction (Mooney, Ryan, Uhing, Reid, & Epstein, 2005). A fourth review suggested SM was comprised of four processes: self-monitoring, self-evaluation, self-reinforcement, and self-instruction (Stage & Quiroz, 1997). The fifth and final review examined the impact of SM on behavioral outcomes by applying the What Works Clearinghouse (WWC) inclusion criteria to the studies collected in the 2009 Briesch & Chafouleas review (Maggin et al., 2013). The fifth review defined SM as including five steps: self-determining a performance goal, self-monitoring, self-recording, self-charting, self-evaluation, and self-reinforcement. To summarize, though researchers define and combine the elements of SM differently, we may observe studies in this review that include all 11 SM elements—though the likelihood of this occurring is low.

Adding to the variety of definitions and typologies that identify a SM intervention, very little research examines the impact of training students in SM. Studies do suggest that the presence of student training has an impact on the success of a SM intervention. For example, existing studies suggest that when students are provided with training in SM they evidence

improved academic (Greiner & Karoly, 1976; Harris, 1986) and behavioral performance outcomes (Harris, Friedlander, Saddler, Frizzelle, & Graham, 2005; Ninness, Fuerst, Rutherford, & Glenn, 1991). However, no studies directly compare training vs. no training or unpack the most effective SM training strategies that appear to enhance outcomes. Regardless, it is a widely-accepted best practice to train participants in the steps of a cognitive-behavioral intervention before they engage in the actual procedures of the intervention (Kazdin, 2003). As such, a novel opportunity for the current review is to capture characteristics of student training in SM to examine the impact of training on outcomes. To examine the impact of training students in SM, a summary of what is currently known about training students in the SM is offered before we describe SM procedures.

The SM student training

Considerations for training students in SM are provided by Cole et al. (2000), Shapiro et al. (2002), Strayhorn (2002b), and Lane et al. (2010)—however very little empirical research identifies which of these considerations are more important than others.

Some research suggests training is more effective when it (a) is sequenced, (b) uses active learning modalities such as modelling and rehearsal, (c) is focused on individual needs, and (d) explicitly defines the behavioral skills required to engage in SM (Durlak et al., 2011). Training is also enhanced when students have opportunities to practice the procedures of SM (Schunk & Zimmerman, 1998; Strayhorn, 2002a, 2002b), and when those practice sessions result in specific and formative feedback aimed at improving accuracy (Shute, 2008). Some research also indicates that feedback should also be supportive when affirming or correcting student SM efforts (Bandura, 1994; Dweck, 1975; Mueller & Dweck, 1998; Zimmerman, 1989). Praise for effort communicates that failure is a part of learning and effort matters more than achievement, effectively lowering the effect of appraisal on performance of a required task (Dweck, 2006). Such messages also engender the development of quality relations between students and teachers (Wentzel, 1991) and encourage children to practice and integrate SM skills (Lewis, 2000).

There are three possibilities when determining the focus of training for a SM intervention. The first possibility is to determine whether the student *can* perform a task. If the student cannot perform the task or achieve the desired outcome, then he or she will require direct instruction in the basic steps to perform that task. If the student has the ability to perform a task but requires assistance with doing the task fluently, smoothly, or with more confidence, a second possibility can be considered. In this instance it must be determined whether the student requires additional instructional supports, more practice, or both. Lastly, if the student has the capacity to perform the task fluently but refuses or is not motivated, then the student may need support or reinforcement to engage in the task. Either way, studies suggest that students may be trained in SM and that training can mitigate any of the three performance barriers listed here (Cole et al., 2000; Lane et al., 2010; Shapiro et al., 2002).

Given the number of cognitive and behavioral skills required to engage in SM, a number of

considerations raised by Cole and colleagues (2000) serve as a helpful guide to illustrate how student training in SM will likely vary across studies:

1. What type of students will SM be used with?
2. What type of outcomes will SM be used with?
3. What type of setting will SM be used in?
4. What type of prompt is suitable for the setting?
5. What type of recording device is suitable?
6. What type of reinforcement should be used?

Type of students

Studies have suggested that SM interventions are feasible and effective at improving behavioral outcomes for males and females, students from a variety of racial and ethnic groups, and students in elementary through high school grades (Briesch & Chafouleas, 2009; Shapiro et al., 2002). Studies have suggested SM interventions are effective for both students without disabilities (Wood, Murdock, Cronin, Dawson, & Kirby, 1998) and for those with a range of disabilities—including autism (Koegel, Koegel, Hurley, & Frea, 1992), developmental delays (O'Connell, Boat, & Warner, 2009), ADHD and learning disabilities (Shimabukuro, Prater, Jenkins, & Edelen-Smith, 1999), and emotional and behavioral disorders (Thompson, 2012; Thompson & Webber, 2010). Some studies have suggested that SM is not effective at improving behavioral functioning for individuals with intellectual disabilities (i.e., mental retardation; Shapiro, 1981).

Type of outcomes

The research underlying the effectiveness of SM suggests the intervention is effective at improving both academic and behavioral outcomes (Briesch & Chafouleas, 2009; Mooney et al., 2005). Regarding academic outcomes, SM has been shown to improve academic performance and rates of work completion and accuracy (Carr & Punzo, 1993; Miller, Miller, Wheeler, & Selinger, 1989; Mooney et al., 2005). With regard to behavioral outcomes, studies suggest SM may be used to improve attention, and reduce impulsivity and externalizing behaviors (e.g., talking out, out of seat). Studies also suggest SM has been used to decrease more common acts of insubordination such as off-task behaviors (Blick & Test, 1987; Dunlap & Fox, 1999; Hallahan & Sapona, 1983; Prater, Joy, Chilman, Temple, & Miller, 1991; Webber, Scheuermann, McCall, & Coleman, 1993) as well as acts associated with aggressive and antisocial behaviors (Bennett & Gibbons, 2000; Todd, Horner, & Sugai, 1999).

Type of setting

A variety of settings in a school may condition a SM intervention. Since students spend the majority of their time in the classroom, a majority of studies examine the effects of SM on classroom behaviors. However, some studies have also examined the effects of SM on the effects of behavior on the playground (Koegel et al., 1992), in gym class (Zimmerman &

Kitsantas, 1996), and in the hallways or during other unstructured transitions (Connell, Carta, Lutz, & Randall, 1993). The type of setting is often a result of the target behavior and contextual factors associated with the behavior. Because the setting or context will vary, studies will vary in the types of behavior and prompts used.

Type of prompt

Many different prompts are used in the self-monitoring phase of a SM intervention. Prompts, broadly speaking, come in two forms: internal and external. Internal, also referred to as a self-prompt, generally requires a student to remind herself to reflect on her own behavior. However, the reliability of internal prompts is not well-established and is questionable given the challenges faced by children with attentional and behavioral issues. As such, many studies rely on external prompts (Cole et al., 2000; Shapiro et al., 2002).

External prompts may take the form of a verbal or nonverbal cue delivered by an adult (e.g., verbal reminder, hand signal) or an electronic device (e.g., a watch, a timer). Some studies have used tape recorders and headphones to prompt students (DiGangi, Maag, & Rutherford Jr, 1991; McDougall & Brady, 1998). More recently, studies are beginning to examine the use of personal digital devices as prompts (i.e., laptops, tablets, personal digital assistants, mobile phones, and digital vibrating devices). One possible benefit offered by personal digital devices is that they address issues of reliability during the self-monitoring phase of a SM intervention. Findings suggest the prompts delivered by these devices do not disrupt other students (e.g., vibrating devices), allow practitioners to vary the self-monitoring schedule to fit an individual student's need, and improve the accuracy of self-observation and self-recording (Amato- Zech, Hoff, & Doepke, 2006). The use of mobile technology in the research is beginning to explore the use of digital devices to integrate external prompts with real time digital data collection of SM observations (Fjeldsoe, Marshall, & Miller, 2009; Gulchak, 2008; Mitchem, Kight, Fitzgerald, Koury, & Boonseng, 2007). Obviously, the type of prompt has implications for reliability of the monitoring element of a SM intervention—although some research suggests SM is effective regardless of the accuracy and reliability of the self-observations (Cole et al., 2000). Regardless, the type of recording device will vary across the studies included in the review and will impact how students are trained in SM.

Type of recording device

The type of recording device refers to the characteristics of the tool used to record SM observations (Cole et al., 2000). Broadly speaking, two characteristics are used to describe a monitoring device—interval frequency and observational response options. To increase the accuracy of monitoring data—and arguably the effectiveness of a SM intervention—the recording device should be simple. The characteristics of the device are likely to vary along several lines as determined by student needs and behaviors.

First, a device should be simple, available, portable, and have utility in multiple school settings (e.g., hallways, classroom, playground, gym class; Cole et al., 2000). Availability and

portability increases the device's utility. However, the target behavior and how the behavior is operationally defined will also condition the utility of the device across multiple school settings. That is, some behaviors are just not relevant to all settings or occur in the presence of certain other factors. Lastly, the cost impacts the choice of device in practice and research. Generally, paper-and-pencil interval recording devices are used in most SM studies because they are easily manipulated, highly accessible, and are low-cost (Lane et al., 2010; Shapiro et al., 2002).

Because studies examine the effects of SM with the full range of school-aged students as well as across students with varying disability types, the age and ability of a child are important considerations for determining the format of a recording device. Some devices may only record the presence of an on- or off-task behavior (Bolstad & Johnson, 1972; Harris, 1986; Harris et al., 2005; Shapiro et al., 2002). Other studies may have specific behavioral goals (Thompson & Webber, 2010). However, most studies use time interval formats with multiple intervals spaced at equal time points.

Tables 1 and 2 provide some selected examples of the broad number of interval SM recording devices. While there are many possible formats for the device, Table 1 exemplifies a device useful with young students or those with learning impairments (i.e., few intervals, responses with pictorial options). The example in Table 2 may be used with older students. In this example, there are multiple target behaviors and multiple recording intervals to address an array of complex and competing behaviors.

TABLE 1: RECORDING EXAMPLE FOR A YOUNG STUDENT

	Before Lunch	After Lunch
Raise Hand	🙂 😐 🙁	🙂 😐 🙁
Stayed in Seat	🙂 😐 🙁	🙂 😐 🙁

What type of reinforcements should be used?

Using contingency reinforcers alongside a SM intervention may improve the success of the intervention. For example, requiring a student to meet his or her predetermined goal to earn a positive reinforcer (e.g., extra time at recess or playing a game with a peer) or a negative reinforcer (e.g., earn a pass on completing an assignment) has been shown to improve outcomes (Glynn, Thomas, & Shee, 1973; Webber et al., 1993). Studies also suggest when SM goals are achievable and the contingencies are provided immediately upon goal attainment, SM appears to be more effective (Lane, Kalberg, Bruhn, Mahoney, & Driscoll, 2008).

In summary, there are many aspects to student training and only a few of which are listed here. However, it is generally agreed that students should be trained in the skills needed for

SM before they actually engage in following SM procedures.

TABLE 2: RECORDING EXAMPLE FOR AN OLDER STUDENT

Time	Completed Work		Accepted Direction		Stayed in Assigned Area	
8:00-8:30	Yes	No	Yes	No	Yes	No
8:31-9:00	Yes	No	Yes	No	Yes	No
9:01-9:30	Yes	No	Yes	No	Yes	No
9:31-10:00	Yes	No	Yes	No	Yes	No
10:01-10:30	Yes	No	Yes	No	Yes	No

The SM procedures

The SM procedures consist of students engaging in one or a combination of the 11 processes listed above that constitute a SM intervention. Although procedural aspects would certainly be taught during the training stage, the SM procedural stage directly refers to the cognitive and behavioral processes a student would be expected to engage in during the actual implementation of a SM intervention. The procedural stage may include any one or a combination of the 11 SM elements listed above (Hallahan & Sapona, 1983; Rutherford Jr, Quinn, & Mathur, 2004; Vaughn, Bos, & Schumm, 2011).

During the self-assessment phase, students may *self-select*, *self-define*, and *self-determine* reasonable performance goals to address a target behavior. Even if students are only self-monitoring on- or off-task behavior, they must first select and define a behavior. Ideally, the behavior will be one that occurs at a frequency or rate that allows it to be observed or detected at regular intervals. That is, if a behavior is low frequency then it is unlikely to register or be observed to a degree that self-monitoring will capture useful information about the behavior. Once a student has identified and defined a problem behavior, a goal may also be set to reduce the problem or increase the performance of a preferred replacement behavior. Using observable and measureable terms (i.e., frequency, duration, and/or severity of the behavior), a student may operationally define a goal using positive language (i.e., I will increase my work completion) or negative language (i.e., I will not tap my pencil on my desk). Though some researchers argue that to be considered a true SM intervention, students should directly participate in each of the 11 SM steps, many studies report a process whereby educators select, define, and set performance goals for students (Briesch & Chafouleas, 2009). Students or teachers may also identify reinforcements at this stage should the student

achieve the goal. Once the student has selected, defined, and set a goal to address a behavior, the student is ready to self-monitor his or her performance.

During the self-monitoring phase, students first *self-prompt* or are externally prompted to self-observe. During the *self-observation*, a student reflects upon his or her performance and discriminates whether he or she displayed the target behavior during the interval. During the *self-recording* process, a student would physically record the observation on a schedule (see Tables 1 and 2) to indicate the presence or absence of the target behavior. Following the self-monitoring phase, a student then may evaluate his or her own progress.

During the self-evaluation of performance, students self-chart their performance by calculating percentages or creating graphic images of the data collected (DiGangi et al., 1991). Using the percentages, graphs, or charts, the student can *self-appraise* or compare his or her results to a predefined goal, prior observational data, using teacher or other third party observations, or combination of those standards (Thompson & Webber, 2010). Using the standards, students can determine whether their performance met the standard and—if relevant—*self-administer* selected reinforcements. The performance comparisons may also be used to develop new performance goals before the process is then iteratively repeated.

In summary, the review will examine SM interventions, a widely-used cognitive behavioral intervention that appears to be effective for academic and behavioral outcomes. Though implementing a SM intervention appears straightforward, there are many variations in practice surrounding training and implementation of a SM intervention. Regardless of these variations, we suggest a SM intervention is best defined as a set of strategies that train students to assess, monitor, and evaluate their own behavioral performance. As such, SM consists of two stages: a training stage and a procedural stage. Though some resources are available to suggest best practices and considerations for training students to engage in SM procedures, no manualized SM programs are available that sequence empirically supported elements of student training, which may improve outcomes. As such, it is expected that the type, quality, and degree of student training will vary greatly across studies of SM interventions. In addition, the SM procedural stage will also vary based upon many contextual and child-specific features.

How the Intervention Might Work

There are several important behavioral principals or mechanisms of change underlying a SM intervention. To begin, behavior change can occur by the very function of engaging in the self-monitoring aspect of a SM intervention (Nelson & Hayes, 1981). Also known as the *reactivity* principle, the simple act of collecting SM data regarding one's own behavioral functioning is thought to alter the behavior itself. For example, and observed in studies of self-regulated learning (Bandura, 2005; Cleary & Zimmerman, 2004), students who self-monitored their performance on a number of math problems were systematically introduced to a heightened awareness of the number of problems answered correctly. Change was hypothesized to occur as a direct result of internal reward mechanisms that influence

behavior change. That is, the simple act of observing and recording one's own performance informs and influences reward centers, which alters motivation and behavior (Shapiro et al., 2002). The reactivity phenomenon has also been observed in a variety of other research areas. For example, weight loss was observed in studies where participants monitored daily caloric intake and types of foods eaten without engaging in dieting interventions (Boutelle & Kirschenbaum, 2012; Butryn, Phelan, Hill, & Wing, 2012). The reactivity phenomena has also been observed in studies of SM interventions with persons who have substance and alcohol abuse disorders (Bien, Miller, & Tonigan, 2006; Simpson, Kivlahan, Bush, & McFall, 2005). Although it may appear, on the surface, that no discernible extrinsic reinforcers are present during the SM procedure—the very act of reflecting on behavior, collecting data on behavior, and using that data to evaluate performance over time is a metacognitive activity that alters the targeted behaviors.

Another mechanism related to the theory of change underlying SM—one closely related to the concept of reactivity, self-management, self-awareness, and intrinsic motivation—is the concept of *perceived autonomy* (Deci, Cascio, & Krusell, 1975; Deci & Ryan, 2011; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2007). In a SM strategy, students engage in an act of "perceived autonomy." That is, students are encouraged to self-assess, self-monitor behaviors, and self-evaluate specific behaviors (Algozzine et al., 2001; Field et al., 1998; Wehmeyer & Schwartz, 1997). Because behavior change is often a "top down" activity that is prepared, planned, and applied to students by teachers, supporting student autonomy through the use of a SM intervention improves perceived ownership and motivation to engage in the intervention, which leads to an increased likelihood of positive outcomes (Lane et al., 2010). Indeed, a variety of studies have shown that when teachers engage in autonomy support strategies (e.g., choice making, goal setting, self-observation, instruction in self-control), participants perform tasks consistently better than tasks where autonomy is not supported (DeCharms, 1984). In short, SM is an autonomy support activity that provides students with choice and, as such, they experience increased levels of perceived autonomy.

Furthermore, because SM is an iterative process, students are provided with increased opportunities to practice skills. Opportunities to practice novel skills leads to an increased sense of *self-efficacy* or *competency* surrounding the completion of a required behavioral task (Eccles, Early, Fraser, Belansky, & McCarthy, 1997; Niemiec & Ryan, 2009). With increased competencies, students are more likely to adopt and integrate those external requirements into their repertoire of internalized skills and values (Gagné, 2003). Furthermore, autonomy support, relevant instruction, and increased opportunities to practice and develop competencies have been shown to improve relationships between students and teachers. That is, autonomy support is an important mediator shown to facilitate healthy and trusting student-teacher relationships (Connell et al., 1993; Cox & Williams, 2008; Hamre & Pianta, 2003; Wentzel, 1993; Wentzel, 2002; Wentzel et al., 2007). Quality relations between students with challenging behaviors and their teachers diminishes challenging behaviors (Wentzel et al., 2007) and makes disciplining students who do engage in challenging behaviors more effective (Hamre & Pianta, 2003).

Why it is Important to do the Review?

Five reviews currently examine the impact of SM on student behavioral or academic outcomes. Four of the reviews examine the impact of SM on behavioral performance and one review examines the impact of SM on academic performance. The reviews are summarized in Table 1 in Appendix A. The first review examined the effects of SM on behavioral outcomes for 817 students across 42 studies, 38 of which were single case designs (Fantuzzo & Polite, 1990). In the second review on the effects of SM for behavioral outcomes, Briesch and Chafouleas (2009) followed the format of the Fantuzzo review to examine the effects of SM for 106 students across 30 studies, 21 of which were single case designs. The fourth review by Maggin and colleagues is omitted from Table 1 because the review utilized the same group of studies collected during the 2009 Briesch and Chafouleas review (Maggin et al., 2013). The major difference between the Briesch and Chafouleas (2012) review and a Maggin et al. (2013) review was that the latter applied WWC standards as inclusion criteria for single case studies. A fifth review examined the effects of SM on academic performance for students with disruptive behaviors (Mooney, et al., 2005). In this review, Mooney and colleagues (2005) located 22 studies examining SM with a total of 78 students. Again, the majority of the studies included were single-case designs. The systematic reviews, taken together, strongly suggest that a SM intervention impacts desirable behavioral and academic outcomes. However, the studies have limitations that prompt the current proposed review.

The most noteworthy limitation of prior SM reviews has to do with the methods used to generate summary effect sizes. Although prior reviews did not benefit from emerging methods to generate standardized summary effects, the methods used in the studies likely overestimate the effects of SM. Though a great deal of debate surrounds the best approach for synthesizing findings from single case designs, all of the prior reviews combined single subject and group studies in one review, combined multiple baseline and intervention phases, and used the "no assumptions" approach for estimating summary effects (Kratochwill et al., 2002). The no assumptions effect size is estimated by subtracting the mean of the baseline from the intervention mean and dividing the result by the baseline standard deviation. Such summary estimates, when not properly accounted for, inflate effect size estimates, evidenced in part by summary effects of SM ranging anywhere from 4.19 to 30.25 (Briesch & Chafouleas, 2009; Fantuzzo & Polite, 1990; Mooney et al., 2005).

The Maggin, Briesch, and Chafouleas (2013) review relied on two complimentary yet limited approaches to estimate summary effects. The first statistic used in both studies was the percent of nonoverlapping data (PND)—a common metric developed for use in single subject studies (Scruggs, Mastropieri, & Casto, 1987). The drawback of PND is that the approach does not account for the autocorrelation inherent in single case studies. Furthermore, PND does not account for baseline trends that may explain improvements observed during the treatment phase. The second statistic used in both studies was a standard mean difference effect generated using ordinary least squares models with fixed effects. Such approaches do not account for the wide heterogeneity observed when (a) combining phases within single

case studies, (b) combining effects across single case studies, or (c) combining single case and group designs in single summary effects. The Mooney et al. (2005) review used a standard mean difference to estimate summary effects. However, Mooney (2005) only included the average of the last three data points in each phase—a practice that has been shown to inflate summary effect sizes (Olive & Smith, 2005).

A second limitation of the prior reviews hinges on the search procedures used in each of the studies. That is, the prior studies relied upon (a) the same search terms and (b) the same two databases (i.e., PsycINFO and ERIC). In addition, prior reviews included no "grey" literature strategies to include effects of published and unpublished sources not commercially controlled. To demonstrate that the proposed review has prepared a comprehensive and broader list of search terms and databases, we conducted an initial search and compared our results with the initial results reported in two of the most recent SM reviews. Our preliminary search used the search terms listed below and focused on the following databases: Academic Search Premier, Dissertation Abstracts International, ERIC, MEDLINE, PsycINFO, Social Service Abstracts, Social Work Abstracts, and Sociological Abstracts. The search yielded a total of 5,830 studies (see Table 2 in Appendix A). By comparison, Fantuzzo and Polite (1990) initially located 987 results and Briesch and Chafouleas (2009) located 794 results. Though exclusion criteria will certainly eliminate a substantial number of the results we identified, we feel that our proposed search strategy (i.e., a broader range of terms and databases) will capture a pool of relevant studies not included by our predecessors. In addition, the preliminary results displayed in Table 2 in Appendix A did not exhaust the full list of proposed databases or include efforts to capture grey literature.

Lastly, prior reviews did not take full advantage of emerging meta-analytic techniques. That is, the reviews did not use random effects models to generate summary estimates, did not present visual funnel or forest plots, the reviews lumped single subject and group designs together in the analyses rather than analyse the studies separately, and the reviews did not report results from Q test or I^2 to index the level of heterogeneity present in final estimates or suggest methods to address any observed between-study variance. Prior reviews also did not test moderation models to examine whether outcomes varied by important features of student training, student characteristics, or examine the impact of SM by subtypes of challenging behavior (i.e., direct and indirect forms of aggression, overt and covert antisocial, and common acts of insubordination). Though two prior reviews did attempt to conduct component analyses of SM and investigate how specific elements were related to effect size estimates (Briesch & Chaffouleas, 2009; Fantuzzo & Polite, 1990), those studies did not take advantage of models that may examine whether student participation in each of the independent SM components impacted outcomes. Because researchers routinely hypothesize that direct student involvement in each SM process would impact the success of the intervention, such analyses would make an important contribution to the current state of research underlying the effects of SM.

OBJECTIVES

The purpose of the review is to inform practice and policy by evaluating the effectiveness of SM interventions designed to reduce challenging classroom behaviors. The following research questions guide this study:

1. How effective are SM interventions at reducing challenging classroom behavior/increasing positive and pro-social classroom behavior?

    a. Does the use of emerging meta-analytic techniques for single case designs impact estimated effect sizes compared to prior reviews?

    b. How do efforts to capture all available studies through the use of comprehensive search procedures impact results?

2. What does the existing body of literature tell us regarding the state of research on SM interventions? Including:

    a. How rigorously has SM been evaluated?

        i.   What types of research designs are most commonly used?

        ii. What are the most common measurement instruments used to assess behavioral change attributed to SM (i.e., observations, standardized instruments)?

        iii. What types of analytical strategies are used to impart statistical conclusions as to the effects of SM strategies?

        iv. What methods are commonly used to report SM intervention fidelity?

        v.   Do studies report measurement reliability characteristics in the studies (i.e., alphas, test-retest correlations, Kappas)?

    b. Do existing studies examine mechanisms underlying the SM behavioral change process (e.g., increases in social competence or self-efficacy, perceived autonomy, reactivity, etc.)?

    c. Do student characteristics moderate the success of SM for behavioral outcomes?

        i.   Are the effects of SM moderated by student sex?

        ii. Are the effects of SM moderated by student race/ethnicity?

        iii. Are the effects of SM moderated by student age/grade?

        iv.   Are the effects of SM moderated by regular/special education?

  d. Do intervention characteristics moderate the success of SM for behavioral outcomes?

      i.   Are the effects of SM moderated by student training?

      ii.   Are the effects of SM moderated by length of exposure?

  e. Do behavioral subtypes (i.e., direct and indirect forms of aggression, overt and covert antisocial behavior, and common acts of insubordination) moderate the success of SM interventions for behavioral outcomes?

  f. Do studies communicate strategies for training students in SM—and if so—do training features (i.e., sequenced skills, active learning modalities, sufficient focus, explicit skills) moderate student outcomes?

  g. Does the inherent variation of student participation in each of the 11 SM elements moderate outcomes?

  h. Do studies of behavioral SM strategies examine and report academic outcomes— and if so, what are the average effects of SM strategies for academic outcomes (i.e., grades, standardized tests and instruments)?

  i. Does the level of program fidelity moderate intervention outcomes?

METHODOLOGY

Criteria for including studies in the review

The review will include two analyses. The first will assess the effect of multiple group studies (i.e., randomized control [RCT] and quasi-experimental designs [QED]). The second analysis will examine the effects of studies using single subject designs (SSD). The following criteria will be applied to determine the inclusion of any study to estimate the effects of SM on classroom behavior.

Types of studies

Criteria for including a study in the review will be guided by the Institute of Education Sciences What Works Clearinghouse (IES-WWC) for studies that *meet evidence standards* and *meet evidence standards with reservations.* For SSD studies, each will be evaluated on a case-by-case basis using the *Institute of Education Sciences What Works Clearinghouse* standards for single case designs (http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19). The standards include the following:

- The independent variable is systematically manipulated in the study.

- Each study outcome is measured systematically over time by more than one assessor, and the study collects interrater agreement on at least 20% of the data points in the baseline and the intervention conditions.

- The study includes at least three phases to demonstrate an intervention effect at different points in time (e.g., reversal, multiple baseline).

- Each phase must have an adequate number of data points.

Types of participants

For inclusion, studies must involve students with challenging behaviors who: are of school age (i.e., 5-21 years); of regular or special education status (e.g., emotionally disturbed, learning disabled, other health impaired, etc.); attend an elementary, middle, or secondary school program (i.e., public, alternative, special education, charter, or private school) and present challenging behaviors.

Types of settings

To be included, studies must be conducted in a school setting, including public, alternative, charter, private, or special education settings.

Types of intervention

The review will include only studies that identify use of a SM intervention, defined as a cognitive behavioral intervention that trains students in a set of techniques necessary to self-assess, self-monitor, and self-evaluate behavioral performance using one or a combination of 11 sub-elements.

Types of outcomes

The review will include only studies that report outcomes assessing challenging classroom behavior. The following types of outcomes are of interest in this review: 1) aggressive (e.g., hitting or name calling and spreading rumours or betrayal, (2) antisocial (e.g., stealing, bullying, lying, cheating), and (3) insubordinate behaviors (e.g., off-task/on-task, noncompliance, withdrawal, refusal to cooperate, decreased off-task, or increased on-task behavior).

We will also examine the effects of behavioral SM on academic outcomes for studies reporting those outcomes (i.e., course grades, assignment grades, standardized testing results, work completion, etc.).

Measures of challenging classroom behavior may include standardized measures of aggression, antisocial, or insubordinate and disrupting behavior. Measures of challenging

classroom student behavior may also be assessed using daily classroom observational data. That is, teacher observation or third party observations of student off- or on-task behavior and alterations in any specific challenging behaviors are likely to be observed.

Geographical context

The review will make efforts to include studies conducted in any country and published in languages other than English; however, studies originating from contexts where differences in educational systems may bias or limit results will be excluded, such as countries with single sex or racially segregated educational systems. The inclusion of non-English studies will also be contingent upon resources and availability of interpretation services.

Time period

The review will include studies that have been conducted since 1 January 1988. The proposed start date includes the time frame covered in three of the most recent SM reviews (i.e., Briesch & Chafouleas, 2009; Maggin et al., 2013; Mooney et al., 2005) and picks up at the time where the oldest review terminated search procedures (Fantuzzo & Polite, 1990). This 25-year period should afford ample evidence to evaluate the more contemporary effects of SM.

Exclusion criteria

Studies of SM strategies with students who do not present challenging behaviors as the main reason for the intervention will not be included in the review (e.g., self-regulated learning). In addition, because SM requires the use of multiple metacognitive strategies, SM studies with participants who exhibit cognitive impairments or intellectual disabilities will be excluded.

Search strategy

All search procedures will be conducted by two persons. To retrieve eligible studies, several search strategies will be utilized in an attempt to identify and retain published and unpublished studies. Hand searches will be conducted with sources that are not yet electronically available (i.e., the most recent publications waiting to be indexed). Broadly speaking, we will search available electronic databases, research registries, conduct steps to capture gray literature, and review the reference lists of the three most recent reviews of SM strategies for studies meeting the inclusion criteria. More specifically, we will systematically search:

1) Electronic databases

    a. EBSCO Academic Search Premier
    b. British Education Index
    c. CBCA Education

d. Proquest Dissertations and Theses Database

e. ERIC

f. Education Full Text

g. Australian Education Index

h. MEDLINE

i. PsycINFO

j. Social Service Abstracts

k. Social Work Abstracts

l. Sociological Abstracts

2) Research registries

a. Cochrane Collaboration Library

b. Database of Abstracts of Reviews of Effectiveness

c. National Technical Information Service

d. System for Information on Grey Literature

3) Grey literature sources

a. Internet searches will be conducted on Google and Google Scholar using the same terms used in database searches and Boolean logic in as far as it is possible—though some accommodations will be required for these resources, as they will likely not allow the implementation of our full search strategy. Search parameters will be limited by .org, .gov, and .edu domains.

b. Authors of prior studies will be contacted in an attempt to obtain un-published research findings.

c. Conference abstracts and proceedings will be reviewed to identify potentially relevant studies. Conference searches will include:

   i. The Society for Research on Educational Effectiveness (https://www.sree.org/pages/conferences/index.php), and the

   ii. American Educational Research Association Repository (http://www.aera.net/EventsMeetings/tabid/10063/Default.aspx.

d. Clearinghouse and Government websites will be reviewed to identify potential sources of relevant data:

   i. The US Department of Education's web site contains reports of funded programs and initiatives: http://www2.ed.gov/about/offices/list/opepd/ppss/reports.html

   ii. The Institution of Education Sciences, What Works Clearinghouse

contains reports of intervention investigations:
http://ies.ed.gov/funding/grantsearch/index.asp

    e.   Feedback from subject experts will be solicited, once the database and grey literature searches are complete, with the intent to identify additional sources.

4) Bibliographies of previous literature reviews and all retrieved studies. The reference lists from all prior reviews noted above and all studies located once the search methods and inclusion criteria are applied will be reviewed. Additional studies will be included if they meet the inclusion criteria. Grey literature and reference list search techniques will be recorded and descriptive approaches reported in the Review.

Search terms and keywords

During the search process, all permutations of search terms will be recorded so final terms or descriptors used to generate the report will be available to readers. The following search terms will be used to define the domain, intervention, setting, and outcomes with each database:

1. Population[1]: (school* OR class* OR child* OR student* OR adoles* OR teen* OR elementary OR "high school" OR "junior high" OR "middle school" OR "K-12" OR kindergarten)

   AND

2. Domain: ("social behavio*" OR "challenging behavio*" OR "disrupt*" OR "classroom behavio*" OR "student behavio*" OR aggression OR hitting OR "name calling" OR "spreading rumo*" OR antisocial OR stealing OR bully* OR lying OR cheat* OR insubordination OR noncompliance OR withdrawal OR impulsivity OR inattention OR "refusal to cooperate" OR "emotionally disturbed" OR "special education" OR "behavior dis*" OR "problem behavior" OR "emotional dis*" OR "emotional impair*" OR ADHD OR off-task OR on-task)

   AND

3. Treatment: ("self-monitor" OR "self-control" OR "self-report" OR "self-regulat*" OR "self-efficacy" OR "self-prompt" OR "self-record*" OR "self-observ*" OR "self-evaluat*" OR "self-manage*" OR "internal reward*" OR intrinsic OR autonomy OR ownership OR "goal setting" OR "choice making")

   AND

---

[1] To be used if search engines do not allow the option of filters for population.

4. Outcomes: ("behavio* competenc*" OR "social competenc*" OR "emotional competenc*" OR "socioemotional skill*" OR "soft skill*" OR "social awareness" OR improve* OR "academic development" OR "academic achievement" OR "academic performance" OR attitud* OR "social development" OR "instructional time" OR "emotional stability" OR grades OR "standardized test*" OR "work completion").

Management of references and document retrieval

References will be managed using EndNote. All citations for relevant studies will be retained in a file named after the database of origin. Two reviewers will independently conduct searches as outlined above.

For all search results generated by a specific database or approach, the total number of studies retrieved will be logged and the citations will be saved in an EndNote file named after the database of origin. For any citations located from hand searches of journals not available in electronic format, the citations will be entered into EndNote manually using a separate file entitled "hand search." Next, using the search function within EndNote, two reviewers will independently execute each step of the search strategy and record the results before comparing the number of citations located via each database, source, or hand search. Next, reviewers will examine citations and delete duplicate citations. Duplicates will be removed and saved in a separate file.

To secure the identified studies, the University of Missouri (MU) Libraries system will be utilized. Studies that are published but not available will be requested via the Interlibrary Loan. If unpublished studies are not available online, the authors will be contacted via email to request copies of the full text. Similar steps will be followed for hand searches of the most recent journals and these citations will be manually entered into EndNote.

Once all available studies meeting the search criteria are located, two reviewers will examine the title and abstract of each retained study. The purpose of this step is to remove any studies failing to meet the inclusion criteria outlined above. Any discrepancies between the two independent reviewers will be submitted to the study review team where a majority rule decision will be made regarding the inclusion/exclusion of the study. Any studies excluded at this stage will be placed in a separate folder titled "omitted studies." For all retained studies, we will then obtain electronic or paper copies to facilitate in-depth screening and coding of relevant study information.

Description of Methods Used in Primary Research

Appendix B includes a brief summary of the prior reviews of SM and the methods of each study. Due to several factors (i.e., the nature of a SM intervention, classroom setting, student behaviors), the methods most often used in studies are single subject designs. Many of those studies, according to reviews, do employ withdrawal designs to strengthen internal validity. Some studies examine the effects of SM using within group designs, and a small proportion

will use multiple non-equivalent QED. To date, only one RCT of a SM intervention exists. Outcomes are generally measured using repeated observations of behavior (i.e., teacher or third party observations) and standardized pre, post, and follow-up ratings of behavior change. As a result, behavior change data from SM studies is generally analyzed by comparing baseline and intervention percentages using *t*-tests and visual or graphical displays of baseline and intervention functioning.

Criteria for Determination of Independent Findings

We are interested in one primary outcome, challenging classroom behavior, and one secondary outcome, academic achievement. All codable effect sizes for these two outcomes will be extracted using an electronic version of a preformatted abstracting form (see Appendix B), and each outcome will be analysed separately. We anticipate that included studies will use multiple measures for each outcome, multiple reports of the same outcome measure, multiple follow-up time points, and possibly more than one counterfactual condition. These circumstances create statistical dependencies that violate assumptions of standard meta-analytic methods. In order to ensure independence of study-level effect sizes, we will include only one effect size estimate from each independent sample on each outcome construct (i.e. challenging behaviors, academic achievement).

For cases in which a study uses multiple measures (i.e., daily observation and a standardized instrument) of the same construct, we will code data for each measure and create a study level average across the measures. Although we are conceptualizing the outcome of challenging behavior as encompassing the subcategories of behaviors noted above (i.e. aggression, antisocial, and insubordinate behavior), we recognize our definition of challenging behaviors encompasses a range of behaviors that may cluster into conceptually distinct outcomes. Therefore, we plan to code the specific subtype of behavior being measured for all codable effect sizes. The strategy will allow an examination of the relation between SM for different challenging behavior subtypes and to conduct sensitivity analyses of the impact of aggregating the subtypes into one study level effect size as well.

In cases of multiple reports on the same outcome (i.e. parent and child report), we will code data for each report and take an average of all reports. Because averaging reports may result in masking different viewpoints of different reporters, we will examine their concordance beforehand by performing sensitivity analysis to see if results vary across reporters (Littell et al., 2008).

In cases where multiple points of follow-up are provided, we will code follow-up points to conduct a separate analysis for effect sizes comparing studies with similar points of follow-up. In cases where studies report multiple measures or multiple reports on the same outcome, the mean of the measures will be taken to estimate a study-level average across the measures of the same construct. However, and reflecting the exploratory nature of the impact of behavioral SM, we will collect information with regards to any academic outcomes reported to examine the effects of behavioral SM on academic performance. In the case of

multiple counterfactual conditions, we will select the comparison condition that is most similar to those in the other included studies.

In single case designs with multiple observations during the baseline and intervention phases, we will extract averages and standard deviations observed from multiple observations during each phase of the design. To reliably capture the information, we propose to use a graphical program such as *UnGraph* (Hedges, Pustejovsky, & Shadish, 2012; Shadish et al., 2009) to extract numeric data from the graphics.

Once numeric data are extracted, we will rely on the statistical guidance provided by Hedges and colleagues (2012) to extract smoothed estimates of functioning for each baseline and treatment phase. To apply the standards offered by Hedges and colleagues, we will collect and code data from single case studies to reflect the number of cases in each study (in cases of multiple cases in one study), the number of phases of treatment alteration, the number of data points per phase, and the relevant outcome values. In studies where there are multiple follow-up observations, we will also extract and code data in a manner that will allow us to examine the effect sizes of SM on student behavior from a longitudinal perspective.

In cases where we examine that a particular study may be reported in multiple reports or in the more likely case where multiple experiments are conducted across several cases but reported in one study, care will be taken to ensure that the study findings are represented only once in the summary effects. If it is unclear whether reports and studies provide independent findings, the authors of the reports will be contacted to clarify the discrepancy.

Details of Study Coding Categories

All studies retained from the search criteria outlined above will be retrieved and data will be coded and extracted using the instrument in Appendix B. Section A of the coding in Appendix B will capture all bibliographic information needed to locate the study (i.e., report type, how study was located, the country and geographic location of origin, and the language of the study). Section A will also capture all information needed to screen and determine the relevance of a particular study. Relevant study characteristics that need to be captured include whether the study type meets selected WWC design standards, whether the study takes place in a school setting, whether the study outcomes examine challenging behavior, etc.

Sections B and C will capture information relevant to the study methods dependent upon study design. Multiple group (RCT and QED) designs will use Section B whereas single case studies (SSD) will use Section C. Both sections also collect information needed to establish the quality of the state of the research underlying SM. Relevant data will include method of assignment, degree of attrition, information needed to estimate effect sizes for continuous and dichotomous outcomes, information related to group equivalence, type and scale of outcome measures, reliability estimates, the analytic strategy used, and the assessment of fidelity. For studies that fail to provide adequate detail to estimate effect size, the review

team will directly contact the study authors to gather the necessary and relevant information.

Section D will capture information regarding participants. The information included here will focus on demographic data of all study participants, including subtypes of challenging behavior (see definition provided above), age and grade, race/ethnicity, sex, and whether the participant received free and reduced lunch. Information regarding the type of setting will also be captured in the Section C.

Section E will capture details specific to the SM intervention itself. The information here will include the number of SM elements that students are exposed to, the name of the SM intervention, information surrounding whether teacher monitoring procedures accompanied student self-monitoring, and whether students compared their data with teacher data or other means, etc. Information also captured in this section will provide some understanding of the impact of training for outcomes, including length of training, the intervention agent providing the training, and features of training modalities used. Lastly, information surrounding the type of goal selected, the types of prompts, and the features of the recording device are captured along with information on whether reinforcements were used.

Lastly, Section F will capture details with regards to fidelity or the extent to which delivery of an intervention adheres to a study's SM protocol. To understand the relationship between fidelity and SM, we want to examine the types of fidelity that are recorded (structural or procedural) and how those measures are collected and reported as well as the levels of fidelity provided in studies.

To increase reliability of coding procedures and decisions, two research team members will independently code 100% of all selected studies. All coded studies will be cross-referenced. Care will be taken to estimate inter-rater reliabilities for the coding scheme by using Cohen's Kappa. Two Kappa measures will serve as a measure of (a) between and (b) within study inter-rater agreement. The two Kappa's will be averaged to communicate a total study inter-rater reliability estimate to be included in our final report. Any emerging discrepancies between the two coders will be submitted to the research team for a final decision using majority rule.

To code studies, we elect to replicate our coding document in Appendix B using Google Docs. Using Google Docs, data entered into an electronic form is directly linked to a spreadsheet (https://support.google.com/drive/answer/87809?hl=en). This strategy will allow reduce the likelihood of data entry errors. Google permits the spreadsheet data to then be saved as a Microsoft Excel spreadsheet, which can then be imported into conventional analytical programs to conduct descriptive and meta-analyses (i.e., analysis of heterogeneity, sensitivity analysis, analysis for publication bias, main study effects, and moderation models) using the extracted data. To conduct meta-analyses, we propose to use SPSS. In cases where some proposed analytical approaches may not be accessible in SPSS, the data can be transferred into Stata (StataCorp, 2005), a general purpose statistical program that has freely accessible macros written by experts in meta-analysis that can be downloaded to

address issues related to publication bias (i.e., funnel plot, rank correlation test, Egger's test, trim and fill). In addition, SPSS macros are available from developers of methods for estimating a standardized mean difference statistic for single case designs that is similar to Cohen's *d* for between study effect size estimates (Hedges et al., 2012).

Statistical Procedures and Conventions

Once the data are imported into SPSS, we will conduct descriptive analyses on all variables of interest to provide information regarding

- the characteristics of study participants (e.g., subtypes of challenging behaviour, education status, gender, race, income level, grade, age),

- settings where SM studies are commonly situated (e.g., school type, classroom type, geographical location, community characteristics),

- relevant intervention characteristics (e.g., defining features and length of student training, intervention agents, number of SM steps involved, nature of goals [reduce challenging behaviors or increase use of replacement behaviors], types of prompts, recording device characteristics [average interval recordings and number of response options], characteristics of reinforcements [positive or negative], average length of exposure to the intervention, etc.).

- the rigor surrounding the current state of the research method underlying the effectiveness of SM (strengths and weaknesses in sample and sampling methods, designs, measures, analysis).

Following descriptive analysis, the data will be examined for the impact of publication bias by examining the symmetry of a funnel plot where the relative effect size and sample size are plotted for each study. Should the funnel plot indicate the possibility of bias, Trim and Fill methods may be one option to adjust for publication bias (Cooper, Hedges, & Valentine, 2009). Stata macros are freely available to assist the authors to adjust for publication bias using the Trim and Fill method (Duval & Tweedie, 2000). Sensitivity analysis will also be conducted using forest plots to examine any biasing effects due to outliers. If necessary, additional sensitivity analysis will be conducted should other issues arise that may impede our confidence in the pooled estimates.

Next, we will estimate effect sizes for each individual study. For RCT and QED studies, we will use the natural log of the odds ratio effect size statistic for dichotomous outcomes and the standardized mean difference with Hedges' *g* correction for continuous outcomes. For SSD studies that will be included using the WWC criteria, we elect to estimate study level effect sizes using a standardized mean difference effect, which similar to a Cohen's *d* as described by Hedges and colleagues (2012). Indeed, prior comparative studies of available effect size options for SSD studies (i.e., PND, general linear approaches, and the percentage

reduction measure) suggest that the standardized mean difference effect is the most consistent effect size metric for single case designs (Olive & Smith, 2005). That said, the authors acknowledge that there are other means for producing study level effect sizes for small datasets that are characteristic of SSD.

In group designs where groups of students, rather than students, have been assigned randomly, it will be necessary to correct for the clustering effects in the calculation of effects. If studies report cluster-adjusted data, the adjusted data will be utilized to conduct data analysis. If cluster adjustment has not been accounted for, corrections will be made using the intraclass correlation coefficient prior to calculating effect size (Littell, Corcoran, & Pillai, 2008).

Following the estimation of individual study level effects, we will then conduct separate meta-analyses to pool studies into summary effects for group (i.e., RCT and QED studies) and SSD studies. For RCT and QED studies, a weighted mean effect will be calculated by weighting each study level effect size by the inverse of its variance (Littell et al., 2008). For SSD studies, we propose to use the d-statistic as described by Hedges and colleagues (2012). Because methods for synthesizing the results across SSD are developing, further research and debate may arise which could provide a compelling case for using a different summary effect statistic for SSD. However, we elect to use the standardized mean difference effect size for single case designs because it is consistent with our study level approach and it is similar to the effect statistic we are using for between group experiments. Furthermore, the standardized mean difference effect size across SSD studies addresses 1) the number of time-points within cases; 2) the number of cases; 3) the autocorrelation arising from repeated measures; 4) the ratio of within case variability to total variability; 5) a Hedges' g correction for the small sample size; and 6) variation in the number of data points phase and repetitions of the AB phases (Hedges et al., 2012). Although methods of calculating effect size estimates for SSD are relatively new and in development, we have selected this technique at this point in time because it is fairly well-developed and we have access to the code and resources to apply the method. Lastly, random effects models are the most appropriate choice for the estimation of effects sizes as we anticipate significant between study heterogeneity when estimating the effects between RCT and QED studies and between the SSD studies.

Following the estimation of summary effects, we will conduct a test of homogeneity (Q-test) to compare the observed variance to what would be expected from sampling error. The I^2 statistic will also be used to describe the percentage of total variation across studies due to the heterogeneity rather than chance. We will also construct a forest plot displaying study-level mean effect sizes and 95% confidence intervals for the included studies to provide opportunity for visual analysis of the precision of the estimated effect sizes, detection of studies with extreme effects, and information regarding heterogeneity of studies.

Following the estimation of main intervention effects and sensitivity analysis, moderation

models will examine whether subgroup differences exist. Provided that ample information exists, we will use moderation models to examine whether the effects of SM differ by student characteristics (e.g., age, gender, grade, race/ethnicity, regular or special education status, socioeconomic status, etc.). Moderation analyses will also be used to examine whether any study characteristics moderate the effects of SM strategies (e.g., design features, standardized versus observational studies, reliability estimates, measurement approaches, etc.). Provided that an ample number of studies and coded data exist, moderation models will then be used to examine whether the presence of student training and whether the features of student training moderate the relationship between SM and outcomes.

No modifications will be made to model missing data in any of the analytical steps outlined above. However, moderation models may be used to determine the differential effect sizes for studies that report characteristics to conduct moderation models and those that do not.

Treatment of Qualitative Research

Although the reviewers will note any qualitative data throughout the search process, we currently have no intention of conducting qualitative analysis as part of this review.

REFERENCES

Algozzine, B., Browder, D., Karvonen, M., Test, D. W., & Wood, W. M. (2001). Effects of interventions to promote self-determination for individuals with disabilities. *Review of Educational Research, 71*(2), 219-277.

Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in the Schools, 43*(2), 211-221.

Bandura, A. (1994). *Self-efficacy*. Wiley Online Library.

Bandura, A. (2005). The primacy of self-regulation in health promotion. *Applied Psychology, 54*(2), 245-254.

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.

Bien, T. H., Miller, W. R., & Tonigan, J. S. (2006). Brief interventions for alcohol problems: a review. *Addiction, 88*(3), 315-335.

Blick, D. W., & Test, D. W. (1987). Effects of self-recording on high-school students' on-task behavior. *Learning Disability Quarterly, 10*(3), 203-213.

Bolstad, O. D., & Johnson, S. M. (1972). Self-regulation in the modification of disruptive classroom behavior. *Journal of Applied Behavior Analysis, 5*(4), 443-454.

Boutelle, K. N., & Kirschenbaum, D. S. (2012). Further support for consistent self-monitoring as a vital component of successful weight control. *Obesity Research, 6*(3), 219-224.

Brauner, C. B., & Stephens, C. B. (2006). Estimating the prevalence of early childhood serious emotional/behavioral disorders: Challenges and recommendations. *Public Health Reports, 121*(3), 303-310.

Bridgeland, J., Bruce, M., & Hariharan, A. (2013). *The missing piece: A report for CASEL: Collaborative for Academic, Soical and Emotional Learning.* Retrieved from http://casel.org/wp-content/uploads/casel-themissingpiece-report.pdf

Briesch, A. M., & Chafouleas, S. M. (2009). Review and analysis of literature on self-management interventions to promote appropriate classroom behaviors (1988–2008). *School Psychology Quarterly, 24*(2), 106-118.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived

self-efficacy in classroom management. *Teaching and Teacher Education, 16*(2), 239-253.

Bushaw, W. J., & Lopez, S. J. (2010). A time for change: The 42nd Annual Phi Delta Kappa/Gallup Poll of the Public's Attitudes Toward the Public Schools. *Phi Delta Kappan, 92*(1), 9-26.

Butryn, M. L., Phelan, S., Hill, J. O., & Wing, R. R. (2012). Consistent self-monitoring of weight: A key component of successful weight loss maintenance. *Obesity, 15*(12), 3091-3096.

Carr, S. C., & Punzo, R. P. (1993). The effects of self-monitoring of academic accuracy and productivity on the performance of students with behavioral disorders. *Behavioral Disorders, 18*(4), 241-250.

Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools, 41*(5), 537-550.

Clunies-Ross, P., Little, E., & Kienhuis, M. (2008). Self-reported and actual use of proactive and reactive classroom management strategies and their relationship with teacher stress and student behaviour. *Educational Psychology, 28*(6), 693-710.

Cole, C. L., Marder, T., & McCann, L. (2000). Self-monitoring. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessments of child and adolescent behavior* (pp. 121-149). New York, NY: Guildford Press.

Connell, M. C., Carta, J. J., Lutz, S., & Randall, C. (1993). Building independence during in-class transitions: Teaching in-class transition skills to preschoolers with developmental delays through choral-response-based self-assessment and contingent praise. *Education & Treatment of Children, 16*(2), 160-174.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *Handbook of research synthesis and meta-analysis*. Russell Sage Foundation.

Cox, A., & Williams, L. (2008). The roles of perceived teacher support, motivational climate, and psychological need satisfaction in students' physical education motivation. *Journal of Sport & Exercise Psychology, 30*(2), 222-239.

DeCharms, R. (1984). Motivation enhancement in educational settings. *Research on Motivation in Learning, 1*, 275-310.

Deci, E. L., & Cascio, W. F. (1972). *Changes in intrinsic motivation as a function of negative feedback and threats*. Paper presented at Eastern Psychological Association Meeting, Boston, MA.

Deci, E. L., Cascio, W. F., & Krusell, J. (1975). Cognitive evaluation theory and some comments on the Calder and Staw critique. *Journal of Personality and Social psychology, 31*(1), 81-85.

Deci, E. L., & Ryan, R. M. (2011). Self-determination theory. In P. A. M. van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 416-433). London: SAGE.

DiGangi, S. A., Maag, J. W., & Rutherford Jr, R. B. (1991). Self-graphing of on-task behavior: Enhancing the reactive effects of self-monitoring on on-task behavior and academic performance. *Learning Disability Quarterly, 14*(3), 221-230.

Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social psychology, 53*(6), 1146--1158.

Dunlap, G., & Fox, L. (1999). A demonstration of behavioral support for young children with autism. *Journal of Positive Behavior Interventions, 1*(2), 77-87.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405-432.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* Random House Digital, Inc.

Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social psychology, 31*(4), 674-685.

Eccles, J. S., Early, D., Fraser, K., Belansky, E., & McCarthy, K. (1997). The relation of connection, regulation, and support for autonomy to adolescents' functioning. *Journal of Adolescent Research, 12*(2), 263-286.

Fantuzzo, J. W., Polite, K., Cook, D. M., & Quinn, G. (1988). An evaluation of the effectiveness of teacher-vs. student-management classroom interventions. *Psychology in the Schools, 25*(2), 154-163.

Field, S., Martin, J., Miller, R., Ward, M., & Wehmeyer, M. (1998). *A Practical Guide for Teaching Self-Determination.* Council Exceptional Children.

Fjeldsoe, B. S., Marshall, A. L., & Miller, Y. D. (2009). Behavior change interventions delivered by mobile telephone short-message service. *American Journal of Preventive Medicine, 36*(2), 165-173.

Gagné, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation and Emotion, 27*(3), 199-223.

Glynn, E. L., Thomas, J. D., & Shee, S. M. (1973). Behavioral self-control of on-task behavior in an elementary classroom. *Journal of Applied Behavior Analysis, 6*(1), 105-113.

Grayson, J. L., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education, 24*(5), 1349-1363.

Greiner, J. M., & Karoly, P. (1976). Effects of self-control training on study activity and academic performance: An analysis of self-monitoring, self-reward, and systematic-planning components. *Journal of Counseling Psychology, 23*(6), 495-502.

Gresham, F. M., Lane, K. L., & Lambros, K. M. (2000). Comorbidity of conduct problems and ADHD: Identification of" fledgling psychopaths". *Journal of Emotional and Behavioral Disorders, 8*(2), 83-93.

Gulchak, D. J. (2008). Using a mobile handheld computer to teach a student with an emotional and behavioral disorder to self-monitor attention. *Education & Treatment of Children, 31*(4), 567-581.

Hallahan, D. P., & Sapona, R. (1983). Self-monitoring of attention with learning-disabled children past research and current issues. *Journal of Learning Disabilities, 16*(10), 616-620.

Hamre, B. K., & Pianta, R. C. (2003). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*(2), 625-638.

Harris, K. R. (1986). Self-monitoring of attentional behavior versus self-monitoring of productivity: Effects on on-task behavior and academic response rate among learning disabled children. *Journal of Applied Behavior Analysis, 19*(4), 417-423.

Harris, K. R., Friedlander, B. D., Saddler, B., Frizzelle, R., & Graham, S. (2005). Self-monitoring of attention versus self-monitoring of academic performance effects among students with ADHD in the general education classroom. *Journal of Special Education, 39*(3), 145-157.

Hastings, R. P., & Bham, M. S. (2003). The relationship between student behaviour patterns and teacher burnout. *School Psychology International, 24*(1), 115-127.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *IZA Discussion Paper Series No. 6580.* Retrieved from http://ftp.iza.org/dp6580.pdf

Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference

effect size for single case designs. *Research Synthesis Methods, 3*(3), 224-239.

Joseph, G. E., & Strain, P. S. (2003). Comprehensive evidence-based social-emotional curricula for young children: An analysis of efficacious adoption potential. *Topics in Early Childhood Special Education, 23*(2), 62-73.

Kaiser, B., & Rasminsky, J. S. (2009). *Challenging behavior in elementary and middle school*. Pearson.

Kazdin, A. E. (2003). Psychotherapy for children and adolescents. *Annual Review of Psychology, 54*(1), 253-276.

Koegel, L. K., Koegel, R. L., Hurley, C., & Frea, W. (1992). Improving social skills and disruptive behavior in children with autism through self-management. *Journal of Applied Behavior Analysis, 25*(2), 341-353.

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

Lane, K. L., Kalberg, J. R., Bruhn, A. L., Mahoney, M. E., & Driscoll, S. A. (2008). Primary prevention programs at the elementary level: Issues of treatment integrity, systematic screening, and reinforcement. *Education and Treatment of Children, 31*(4), 465-494.

Lane, K. L., Menzies, H. M., Bruhn, A. L., & Crnobori, M. (2010). *Managing challenging behaviors in schools: Research-based strategies that work*. Guilford Press.

Leff, S. S., & Crick, N. R. (2010). Interventions for relational aggression: Innovative programming and next steps in research and practice. *School Psychology Review, 39*(4), 504-507.

Lewis, J. M. (2000). Repairing the bond in important relationships: A dynamic for personality maturation. *American Journal of Psychiatry, 157*(9), 1375-1378.

Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*. Oxford: Oxford University Press.

Maggin, D. M., Briesch, A. M., & Chafouleas, S. M. (2013). An application of the What Works Clearinghouse standards for evaluating single-subject research synthesis of the self-management literature base. *Remedial and Special Education, 34*(1), 44-58.

McDougall, D., & Brady, M. P. (1998). Initiating and fading self-management interventions

to increase math fluency in general education classes. *Exceptional Children, 64*, 151-166.

Miller, M., Miller, S. R., Wheeler, J., & Selinger, J. (1989). Can a single-classroom treatment approach change academic performance and behavioral characteristics in severely behaviorally disordered adolescents: An experimental inquiry. *Behavioral Disorders, 14*(4), 215-225.

Mitchem, K. J., Kight, J., Fitzgerald, G., Koury, K. A., & Boonseng, T. (2007). Electronic performance support systems: An assistive technology tool for secondary students with mild disabilities. *Journal of Special Education Technology, 22*(2), 1-14.

Mooney, P., Ryan, J. B., Uhing, B. M., Reid, R., & Epstein, M. H. (2005). A review of self-management interventions targeting academic outcomes for students with emotional and behavioral disorders. *Journal of Behavioral Education, 14*(3), 203-221.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social psychology, 75*, 33-52.

Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children, 71*(1), 59-73.

Nelson, R. O., & Hayes, S. C. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification, 5*(1), 3-14.

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education, 7*(2), 133-144.

Ninness, H. A., Fuerst, J., Rutherford, R. D., & Glenn, S. S. (1991). Effects of self-management training and reinforcement on the transfer of improved conduct in the absence of supervision. *Journal of Applied Behavior Analysis, 24*(3), 499-508.

O'Connell, M. E., Boat, T., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. National Academies Press.

Olive, M. L., & Smith, B. W. (2005). Effect size calculations and single subject designs. *Educational Psychology, 25*(2-3), 313-324.

Parke, R. D., & Slaby, R. G. (1983). The development of aggression. In W. Damon (Ed.) *Handbook of child psychology, Volume 4*, (pp. 547-641). Wiley.

Prater, M. A., Joy, R., Chilman, B., Temple, J., & Miller, S. R. (1991). Self-monitoring of on-

task behavior by adolescents with learning disabilities. *Learning Disability Quarterly, 14*(3), 164-177.

Rothbart, M. K., & Rueda, M. R. (2005). The development of effortful control. In U. Mayr, E. Awh, S. W. Keele, & M. I. Posner (Eds.) *Developing individuality in the human brain: A tribute to Michael I. Posner* (pp. 167-188). American Psychological Association.

Rutherford Jr, R. B., Quinn, M. M., & Mathur, S. R. (2004). *Handbook of research in emotional and behavioral disorders.* Guilford Press.

Satcher, D. (2004). School-based mental health services. *Pediatrics, 113*(6), 1839-1845.

Schunk, D. H., & Zimmerman, B. (1998). *Self-regulated learning: From teaching to self-reflective practice.* Guilford Press.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24-33.

Shadish, W. R., Brasil, I. C., Illingworth, D. A., White, K. D., Galindo, R., Nagler, E. D., & Rindskopf, D. M. (2009). Using UnGraph to extract data from image files: Verification of reliability and validity. *Behavior Research Methods, 41*(1), 177-183.

Shapiro, E., Durnan, S., Post, E., & Skibitsky-Levinson, T. (2002). Self-monitoring procedures for children and adolescents. In M. R. Shinn, H. M. Walker, & G. Stoner *Interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 433-454). NASP Publications.

Shapiro, E. S. (1981). Self-control procedures with the mentally retarded. *Progress in Behavior Modification, 12*, 265-297.

Shapiro, E. S., & Cole, C. L. (1994). *Behavior change in the classroom: Self-management interventions.* Guilford Press.

Shimabukuro, S. M., Prater, M. A., Jenkins, A., & Edelen-Smith, P. (1999). The effects of self-monitoring of academic performance on students with learning disabilities and ADD/ADHD. *Education and Treatment of Children, 22*(4), 397-414.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Simpson, T. L., Kivlahan, D. R., Bush, K. R., & McFall, M. E. (2005). Telephone self-monitoring among alcohol use disorder patients in early recovery: A randomized study of feasibility and measurement reactivity. *Drug and Alcohol Dependence,*

79(2), 241-250.

Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review, 26*(3), 333-368.

StataCorp, L. (2005). Stata statistical software: Release 9. *College Station.*

Strayhorn, J. M. (2002a). Self-control: Theory and research. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(1), 7-16.

Strayhorn, J. M. (2002b). Self-control: Toward systematic training programs. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(1), 17-27.

Thompson. A. M. (2012). *A randomized trial of the Self-management Training And Regulation Strategy (STARS): A selective intervention for students with disruptive classroom behaviors.* Ann Arbor, MI: Proquest Publishers.

Thompson, A. M., & Webber, K. C. (2010). Realigning student and teacher perceptions of school rules: A behavior management strategy for students with challenging behaviors. *Children & Schools, 32*(2), 71-79.

Todd, A. W., Horner, R. H., & Sugai, G. (1999). Self-monitoring and self-recruited praise: Effects on problem behavior, academic engagement, and work completion in a typical classroom. *Journal of Positive Behavior Interventions, 1*(2), 66-122.

USDOE. (2005). Special analysis 2005-mobility in the teacher workforce. Retrieved from from http://nces.ed.gov/pubs2005/2005094_Analysis.pdf

USDOE. (2006). School survey on crime and safety (SSOCS). Retrieved from http://nces.ed.gov/surveys/ssocs/tables/el_2006_tab_09.asp

Vaughn, S., Bos, C. S., & Schumm, J. S. (2011). *Teaching students who are exceptional, diverse, and at risk in the general education classroom.* Pearson.

Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices.* Wadsworth Publishing Company.

Webber, J., Scheuermann, B., McCall, C., & Coleman, M. (1993). Research on self-monitoring as a behavior management technique in special education classrooms: A descriptive review. *Remedial and Special Education, 14*(2), 38-56.

Wehmeyer, M., & Schwartz, M. (1997). Self-determination and positive adult outcomes: A follow-up study of youth with mental retardation or learning disabilities. *Exceptional Children, 63*, 245-255.

Wentzel, K. R. (1991). Social competence at school: Relation between social responsibility and academic achievement. *Review of Educational Research, 61*(1), 1-24. doi: 10.3102/00346543061001001

Wentzel, K. R. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology, 85*(2), 357-364.

Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development, 73*(1), 287-301.

Wentzel, K. R., Filisetti, L., & Looney, L. (2007). Adolescent prosocial behavior: The role of self-processes and contextual cues. *Child Development, 78*(3), 895-910.

Wigfield, A., Eccles, J. S., Roeser, R., & Schiefele, U. (2008). Development of achievement motivation. In W. Damon, R. M. Lerner, & D. Kuhn (Eds.) *Child and adolescent development: An advanced course* (pp. 406-434). Wiley.

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2007). *Development of achievement motivation.* Wiley Online Library.

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*, S130-S143.

Wood, S. J., Murdock, J. Y., Cronin, M. E., Dawson, N. M., & Kirby, P. C. (1998). Effects of self-monitoring on on-task behaviors of at-risk middle school students. *Journal of Behavioral Education, 8*(2), 263-279.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329-339.

Zimmerman, B. J., & Kitsantas, A. (1996). Self-regulated learning of a motoric skill: The role of goal setting and self-monitoring. *Journal of Applied Sport Psychology, 8*(1), 60-75.

REVIEW AUTHORS

Lead reviewer:

Name:	**Aaron M. Thompson, PhD MSW**
Title:	Assistant Professor
Affiliation:	University of Missouri
Address:	718 Clark Hall, School of Social Work
City, State, Province or County:	Columbia, Missouri
Postal Code:	65211-4470
Country:	United States
Phone:	573.882.0124
Email:	thompsonaa@missouri.edu

Co-authors:

Name:	**Lindsay R. Ruhr, MSW, MPPA**
Title:	Doctoral Student
Affiliation:	University of Missouri
Address:	729 Clark Hall, School of Social Work
City, State, Province or County:	Columbia, Missouri
Postal Code:	65211-4470
Country:	United States
Phone:	573.882.0124
Email:	lrrvw5@mail.missouri.edu

Name:	**Brandy R. Maynard, PhD, LMSW**
Title:	Assistant Professor
Affiliation:	Saint Louis University
Address:	Tegeler Hall, 3550 Lindell Boulevard School of Social Work
City, State, Province or County:	St. Louis, MO
Postal Code:	63103
Country:	United States
Phone:	512-232-2774
Email:	bmaynard@austin.utexas.edu

Name:	**Natasha K. Bowen, PhD MSW**
Title:	Associate Professor
Affiliation:	University of North Carolina at Chapel Hill, School of Social Work
Address:	325 Pittsboro St.
City, State, Province or County:	Chapel Hill, NC
Postal Code:	27599-3550
Country:	United States
Phone:	919.843.2434
Email:	nbowen@email.unc.edu

Name:	**Michael Pelts, MSW**
Title:	Doctoral Student
Affiliation:	University of Missouri at Columbia
Address:	729 Clark Hall, School of Social Work
City, State, Province or County:	Columbia, Missouri
Postal Code:	65211-4470
Country:	United States
Phone:	573.882.0124
Email:	mdpp74@mail.missouri.edu

ROLES AND RESPONSIBLIITIES

- Content: Aaron M. Thompson

- Systematic review methods: Brandy R. Maynard

- Statistical analysis: Statistical analysis will be conducted by Aaron Thompson and Lindsay Ruhr with consultation from Natasha Bowen and Brandy R. Maynard.

- Information retrieval: Aaron M. Thompson and Lindsay Ruhr, in consultation with the University of Missouri's social sciences librarian and Brandy R. Maynard, will conduct all search procedures to be outlined in the search protocol. Natasha K. Bowen and Brandy R. Maynard will resolve discrepancies for study inclusion criteria using the study protocol.

Aaron M. Thompson has worked as a school-based practitioner, educator, and public school administrator. Thompson has published studies and guides on self-management strategies in school settings, including several single subject designs, a randomized

control design of a manualized self-management intervention, and several book chapters on using self-management strategies in school settings. Thompson has also published a review of classroom-based interventions for children with challenging behaviors.

Lindsay R. Ruhr has a master's degree in social work in addition to a master's degree in public policy administration. She is currently a University of Missouri doctoral student. Lindsay has experience working as a program evaluation coordinator for an after-school program serving at-risk youth.

Brandy R. Maynard is currently a postdoctoral fellow at the Meadows Center for Preventing Educational Risk in Austin, Texas, and an Assistant Research Professor at Saint Louis University, and is on the editorial board of the Education Coordinating Group of the Campbell Collaboration. Brandy recently completed a Campbell systematic review on the effectiveness of indicated truancy interventions and has a review protocol under review with the Campbell's Education Coordinating Group to examine the effects interventions for school refusal behavior. Brandy is experienced in systematic review methods and meta-analytic strategies.

Natasha K. Bowen has extensive experience with assessment and evaluation of school-based programs. She is a primary developer of the Elementary School Success Profile Model of Assessment and Prevention, a social-environmental assessment and intervention system to improve the context of school settings.

Michael D. Pelts has a master's degree in social work and is currently a University of Missouri doctoral student. Michael has a range of practice and program evaluation and research related experience with interventions that aim to improve the resiliency of youth who are considered at-risk.

Thompson, Maynard, Ruhr, and Bowen are experienced in research and information retrieval, information synthesis using a variety of statistical programs (e.g., Stata, SAS, MPLUS, SPSS) and methodologies (e.g., Regression, Latent Variable, Latent Variable Growth Curve, Multi-level, and Person-Centered Models).

All team members will contribute with generating a final report to disseminate the findings.

SOURCES OF SUPPORT

 $1,500.00 of internal funding has been dedicated as a research start up fund to support the review and any purchasing costs that may be incurred. In addition, a one-half appointment of a PhD student (50%: $8,000.00) will be dedicated to the timely completion of the project for the 2013-2014 academic calendar. Assistance may be solicited from Campbell with any aspect of the study—particularly with regards to the treatment of single case designs and with translational services.

REQUEST SUPPORT

No support is requested at this time.

DECLARATIONS OF INTEREST

There are no known conflicts of interest.

PRELIMINARY TIME FRAME

 September 2013 – September 2014

Date you plan to submit a draft review: Six months following the approval of the protocol.

PLANS FOR UPDATING THE REVIEW

The lead reviewer will be responsible for updating the review approximately every 3-5 years.

AUTHOR DECLARATION

Authors' responsibilities

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Campbell Collaboration policy. The Campbell Collaboration will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the relevant Coordinating Group within two years of protocol publication. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every three years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

Publication in the Campbell Library

The support of the Campbell Collaboration and the relevant Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review and subsequent updates in the Campbell Library. Concurrent publication in other journals is encouraged. However, a Campbell systematic review should be published either before, or at the same time as, its publication in other journals. Authors should not publish Campbell reviews in journals before they are ready for publication in the Campbell Library. Authors should remember to include the statement: "This is a version of a Campbell review, which is available in The Campbell Library" when publishing in journals or other venues.

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by: Aaron M. Thompson

Date: 10-24-13

APPENDIX A: SUMMARY OF PRIOR REVIEWS

Table 1: Summary of Prior Reviews			
Authors	Fantuzzo & Polite (1990)	Briesch & Chafouleas (2009)	Mooney et al. (2005)
Title of Study	School-Based, Behavioral Self-Management: A Review and Analysis	Review and Analysis of Literature on Self-Management Interventions to Promote Appropriate Classroom Behaviors	A Review of Self-Management Interventions Targeting Academic Outcomes for Students with Emotional and Behavioral Disorders
Publish Source	*School Psychology Quarterly*	*School Psychology Quarterly*	*Journal of Behavioral Education*
Period Reviewed	1967- March 1988	April 1988- March 2008	1970-2002 (1981-2000 included)
Databases	PsycINFO, ERIC	PsycINFO, ERIC	PsycINFO, ERIC
Search criteria and keywords	**4 major content areas** - behavior treatment and behavior modifications, curriculum programs and teaching methods, academic learning and achievement and classroom dynamics and student adjustment. **11 Keywords-** behavioral interventions, classroom behavior, classroom behavior modification, classroom discipline, self-management, self-monitoring, self-regulation, self-reinforcement, self-control, self-instruction, self-report (articles focusing on age groups other than school-aged children were suppressed)	same as Fantuzzo	behavior disorders, emotional disturbance and conduct disorder in combination with academic status, reading, math, science, social studies, testing, academics, special education, self monitoring, self instruction, goal setting, self evaluation, self management, self reinforcement, self regulated learning, strategy instruction
other searches	none specified	Hand search of *Behavior Disorders, Journal of Applied Behavior Analysis, Psychology in the Schools, School Psychology Quarterly, School Psychology Review.*	Hand search of the *Journal of Special Education, Journal of Emotional and Behavioral Disorders, Behavioral Disorders Exceptional Children.* Ancestral search of references from studies.
No. of studies	42	30	22
# of students	817	106	78
Study Designs	Group Comparison(2), Single case (22)	All single subject, not clear	Group (2), single subject
Training	Not addressed specifically in analysis - see p 181, par 3	research team or teachers trained students	not addressed specifically in analysis - see p 204, par 2
Major Findings	Results suggest that the higher the degree of SM the greater improvement in academic performance. Results are limited due to issues with computing effect size.	SM is an effective intervention with broad utility. Generalizations are limited due to reporting and recording differences	A large variety of procedures used. SM produced positive effects on academic outcomes. Evidence supports claim of generalization of findings.

Table 2: Summary of Preliminary Search Results (Number of Studies)			
Source	Briesch & Chafouleas (2009)	Fantuzzo & Polite (1990)	Thompson et al.
Terms	See Table 1	See Table 1	See Below
Academic Search Premier			1,478
Dissertation Abstracts International			399
ERIC	**	**	1,099
MEDLINE			515
PsycINFO	**	**	2,218
Social Service Abstracts			23
Social Work Abstracts			13
Sociological Abstracts			85
Total	794	987	5,830

1. *Population* (school* OR class* OR child* OR student* OR adoles* OR teen* OR elementary OR "high school" OR "junior high" OR "middle school" OR "K-12" OR kindergarten) AND
2. *Domain* ("social behavio*" OR "challenging behavio*" OR "disrupt*" OR "classroom behavio*" OR "student behavio*" OR aggression OR hitting OR "name calling" OR "spreading rumo*" OR antisocial OR stealing OR bully* OR lying OR cheat* OR insubordination OR noncompliance OR withdrawal OR impulsivity OR inattention OR "refusal to cooperate" OR "emotionally disturbed" OR "special education" OR "behavior dis*" OR "problem behavior" OR "emotional dis*" OR "emotional impair*" OR ADHD OR off-task OR on-task) AND
3. *Treatment* ("self-monitor" OR "self-control" OR "self-report" OR "self-regulat*" OR "self-efficacy" OR "self-prompt" OR "self-record*" OR "self-observ*" OR "self-evaluat*" OR "self-manage*" OR "internal reward*" OR intrinsic OR autonomy OR ownership OR "goal setting" OR "choice making") AND
4. *Outcomes* ("behavio* competenc*" OR "social compentenc*" OR "emotional competenc*" OR "socioemotional skill*" OR "soft skill*" OR "social awareness" OR improve* OR "academic development" OR "academic achievement" OR "academic performance" OR attitud* OR "social development" OR "instructional time" OR "emotional stability" OR grades OR "standardized test*" OR "work completion").

Self-Management Intervention Review Data Coding Form

Study ID#: _____ Coder: _____ APA Citation:

Section A – Source Descriptors & Relevance Screening

A1. Report Type   [type]
- ❑ 1. Journal Article
- ❑ 2. Book/book chapter
- ❑ 3. Gov't report (local, state, federal)
- ❑ 4. Conference proceedings
- ❑ 5. Thesis or Dissertation
- ❑ 6. Unpub report (non-gov't, tech report)
- ❑ 7. Other (specify): _____

A2. How study was located
- ❑ 1. Electronic Database
- ❑ 2. Research Registry
- ❑ 3. Grey Literature
- ❑ 4. Hand Search
- ❑ 5. Expert Referral

A3. Country
- ❑ 1. USA   [coun]
- ❑ 2. UK
- ❑ 3. Canada
- ❑ 4. Australia
- ❑ 5. Other (specify): _____

A4. Setting or region
- ❑ 1. Rural   [area]
- ❑ 2. Suburban
- ❑ 3. Urban
- ❑ 4. More than one of these settings

A5. Language if other than English _____   [lang]

A6. Is the study a Self-Management intervention?   [sm]
- ❑ 1. Yes
- ❑ 2. No

A6. If the study is a SSD, does the study design meet IES WWC standards?   [wwc]
- ❑ 1. Yes
- ❑ 2. No

A7. Does the study take place in a school setting?   [schst]
- ❑ 1. Yes
- ❑ 2. No

A8. Does the study examine challenging behavior as an outcome?   [cbeh]
- ❑ 1. Yes
- ❑ 2. No

A9. Type of study   [design]
- ❑ 1. RCT
- ❑ 2. QED
- ❑ 3. SSD-case study
- ❑ 4. SSD-within group

A10. Does the study meet inclusion criteria?   [incl]
- ❑ 1. Yes
- ❑ 2. If No, why: _____

_____

Section B1—Multiple Group (RCT or QED)
Study Methods and Quality Determination

B1.1 Method of assignment to condition(s)   [grp_assign]
- ❑ 1. Random after matching, stratification, blocking, etc.
- ❑ 2. Random, simple
- ❑ 3. Quasi-random-assigned by some naturally occurring process
- ❑ 4. Matched or statistically controlled on pretest measures
- ❑ 5. Does not apply (SSD)
- ❑ 99. Not specified

B1.2 Did the study have attrition greater than 20%?   [grp_attrit]
- ❑ 1. Yes
- ❑ 2. No
- ❑ 99. Not specified

B1.3 Was there differential attrition in the study?
- ❑ 1. Yes
- ❑ 2. No
- ❑ 99. Not specified

B1.4 If matching was used, how were groups matched?   [grp_match]
- ❑ 1. Matched on pretest measure
- ❑ 2. Matched on demographics
- ❑ 3. Matched on both of the above
- ❑ 4. Propensity Score Matching
- ❑ 5. Other matching technique: _____
- ❑ 6. Not enough information to determine
- ❑ 99. Not specified

B1.5 Was group equivalency at pretest established?   [grp_pre]
- ❑ 1. Yes
- ❑ 2. No
- ❑ 99. Not specified

B1.6 Results of statistical comparisons of pretest differences   [grp_pre]
- ❑ 1. No comparisons made
- ❑ 2. No statistically significant differences
- ❑ 3. Significant differences judged unimportant by coder
- ❑ 4. Significant differences judged of uncertain importance by coder
- ❑ 5. Significant differences judged important by coder

B1.7 If groups were non-equivalent, were statistical controls used?   [grp_ctrl]
- ❑ 1. Yes
- ❑ 2. No

B1.8 What type of measure was used to assess behavior change?   [grp_mstyp]
- ❑ 1. Standardized instrument (specify)_____
- ❑ 2. Teacher observation
- ❑ 3. Researcher Observation
- ❑ 4. Self-report
- ❑ 5. More than one of these methods
- ❑ 6. Other _____
- ❑ 99. Not specified

B1.9 Is reliability for Outcome assessment reported?   [grp_rel]
- ❑ 1. Yes
     -Type _____& Level of reliability_____
- ❑ 2. No
- ❑ 99. Not specified

B1.10 What was the analytical strategy used to assess behavior change?
_____
_____

Section B2—Multiple Group (RCT, QED)
Dependent Variables and Effect Size Information

B2.1 For RCT or QED designs with continuous dependent outcomes
    TX group
      ❑ 1. N _____   [cdv_tx_n]
      ❑ 2. Pre-test group mean _____   [cdv_tx_prx]
      ❑ 3. Pre-test standard deviation _____   [cdv_tx_sd]
      ❑ 4. Post-test group mean_____   [cdv_tx_ptx]
      ❑ 5. Post-test standard deviation_____   [cdv_tx_sd]
    Comparison Group
      ❑ 5. N _____   [cdv_ct_n]
      ❑ 6. Pre-test mean _____   [cdv_ct_prx]
      ❑ 7. Pre-test standard deviation _____   [cdv_ct_sd]
      ❑ 8. Post-test group mean_____   [cdv_ct_ptx]
    Test for Comparing Means & Effect Sizes
      ❑ 9. Are the means adjusted in the analysis?   [cdv_adj]
           ❑ 1. Yes (clarify) _____
           ❑ 2. No
      ❑ 10. Statistical value (*t*-test, *F*-value) _____   [cdv_stat]
      ❑ 12. Degrees of freedom _____   [cdv_df]
      ❑ 13. Reported effect size_____   [cdv_es]

B2.2 For RCT or QED designs with dichotomous outcomes
    TX group
      ❑ 1. N _____   [ddv_tx_n]
      ❑ 2. Percent and/or *n* experiencing success _____   ddv_tx_ps]
      ❑ 3. Percent and/or *n* not experiencing success _____ [ddv_tx_pns]
    Comparison Group
      ❑ 5. N _____   [ddv_ct_n]
      ❑ 6. Percent and/or *n* experiencing success _____ [ddv_ct_prx]
      ❑ 7. Percent and/or *n* not experiencing success _____ [ddv_ct_sd]
    Test for Comparing Means & Effect Sizes
      ❑ 8. Statistical value (*chi2*-test, *F*-value) _____ [ddv_stat]
      ❑ 9. Degrees of freedom _____   [ddv_df]
      ❑ 10. Reported effect size_____   [ddv_es]

Section C1—Single Subject (SSD)
Study Methods and Quality Determination

C1.1 The outcomes were measured by more than one assessor   [ssd_assess]
      ❑ 1. Yes, number of assessors_____
      ❑ 2. No
      ❑ 99. Not specified
C1.2 The assessors collected interrater agreement in each phase for 20% of observations
    (Kappa).   [ssd_20p]
      ❑ 1. Yes
      ❑ 2. No
      ❑ 99. Not specified
C1.3 Interrater reliability was =/> .80 for each phase.   [ssd_rel]

❑ 1. Yes
❑ 2. No
❑ 99. Not specified

C1.4 The study included more than one phase.   [ssd_phase]
     ❑ 1. Yes
     ❑ 2. No
     ❑ 99. Not specified

C1.5 Each phase included at least three observations   [ssd_phobs]
     ❑ 1. Yes
     ❑ 2. No
     ❑ 99. Not specified

C1.6 The IV was systematically manipulated by the researcher(s) [ssd_ivman]
     ❑ 1. Yes
     ❑ 2. No
     ❑ 99. Not specified

C1.7 What was the analytical strategy used to assess behavior change?
_____
_____

Section C2—Single Subject (SSD)
Dependent Variable and Effect Size Calculation

C2.1. What type of measure was used to assess behavior change?   [ssd_mstyp]
     ❑ 1. Standardized instrument (specify)_____
     ❑ 2. Teacher observation
     ❑ 3. Researcher Observation
     ❑ 4. Self-report
     ❑ 5. More than one of these methods
     ❑ 6. Other _____
     ❑ 99. Not specified

C2.2. Is reliability for Outcome assessment reported?   [ssd_rel]
     ❑ 1. Yes
         -Type _____& Level of reliability_____
     ❑ 2. No
     ❑ 99. Not specified

C2.3. What was the analytical strategy used to assess behavior change?
_____
_____

C2.4 For SSD-within group designs with continuous outcomes
     ❑ 1. N _____   [ssd_n]
     ❑ 2. Pre-test mean _____   [ssd_pre]
     ❑ 3. Pre-test standard deviation _____   [ssd_presd]
     ❑ 4. Post-test mean_____   [ssd_pt]
     ❑ 5. Post-test standard deviation_____   [ssd_ptsd]
    Test for Comparing Means & Effect Sizes
     ❑ 6. Statistical value (*t*-test, *F*-value) _____   [ssd_stat]
     ❑ 7. Degrees of freedom _____   [ssd_df]
     ❑ 8. Reported effect size_____   [ssd_es]

C2.5. For SSD with dichotomous outcomes
    TX group
     ❑ 1. N _____   [ssd_tx_n]
     ❑ 2. Percent and/or *n* experiencing success _____   [ssd_tx_ps]
     ❑ 3. Percent and/or *n* not experiencing success _____ [ssd_tx_pns]

Comparison Group
- ❑ 5. N _____ [ssd_ct_n]
- ❑ 6. Percent and/or *n* experiencing success _____ [ssd_ct_prx]
- ❑ 7. Percent and/or *n* not experiencing success _____ [ssd_ct_sd]

Test for Comparing Means & Effect Sizes
- ❑ 8. Statistical value (*chi2*-test, *F*-value) _____ [ssd_stat]
- ❑ 9. Degrees of freedom _____ [ssd_df]
- ❑ 10. Reported effect size_____ [ssd_es]

C2.6. For SSD with multiple baselines,
- ❑ 1. Case Id_____ [ssd_ID]
- ❑ 2. Number of phase (A1, B1, A2, B2, etc.) [ssd_ph]
- ❑ 3. Number of obs/phase_____ [ss_obs/ses]

D—Participants, Intervention Agents, and Setting Descriptors

D1. Average Age of participants _____. [age]
D2. Average grade of participants_____. [grd]
D3. Race/Ethnicity [raceth]
- ❑ 1. African American ____%
- ❑ 2. Asian American____%
- ❑ 3. European American ____%
- ❑ 4. Hispanic American____%
- ❑ 5. Other ____%
- ❑ 99. Not specified

D4. Sex [sex]
- ❑ Male____%
- ❑ Female____%

D5. Free or Reduced Lunch [frl]
- ❑ 1. Receiving ____%
- ❑ 2. Not Specified

D7. Setting [schtyp]
- ❑ 1. School
- ❑ 2. Public
- ❑ 3. Private
- ❑ 4. Alternative
- ❑ 5. Charter
- ❑ 6. Other (specify): _____
- ❑ 99. Not specified

D8. School [sch_cntr]
- ❑ 1. Elementary
- ❑ 2. Middle or Junior
- ❑ 3. Secondary
- ❑ 4. Mixture_____
- ❑ 99. Not specified

D9. Primary education status/eligibility [edstat]
- ❑ 1. Regular
- ❑ 2. 504 Plan
- ❑ 3. Special
  - o 4. LD
  - o 5. ED
  - o 6. Other Health Impaired
  - o 7. Other
  - o 8. Mixture (cat. & %)_____
- ❑ 99. Not specified

D10. Who provided the services? [intagt]

❑ 1. Non-School Master's or PhD clinician
❑ 2. School Clinician (Social Worker, Psychologist, Counselor)
❑ 3. Teacher
❑ 4. Other school personnel
❑ 5. Researchers (efficacy trial)
❑ 6. Multiple providers _____(list)
❑ 7. Other: _____(list)
❑ 99. Not specified

D11. Did the provider receive special training on the intervention?   [tr_intagt]
    ❑ 1. Yes
    ❑ 2. No
    ❑ 99. Not Specified

D12. Subtypes of challenging behavior   [behstyp]
    ❑ 1. Aggression (direct or indirect physical or social aggression)
    ❑ 2. Antisocial (overt or covert)
    ❑ 3. Insubordination (noncompliance, withdrawal, refusal, off-task)
    ❑ 99. Not specified

Section E - Intervention Descriptors

E1. Name of intervention: _____

E2. Did students engage in self-assessment?   [int_sa]
    ❑ 1. Self-select target behavior
    ❑ 2. Self-define target behavior
    ❑ 3. Self-determine performance goal[s]
    ❑ 4. Self-identify reinforcers

E3. Did students engage in self-monitoring?   [int_sm]
    ❑ 1. Self-prompt reflect on target behavior
    ❑ 2. Self-observe target behavior
    ❑ 3. Self-record the observation

E4. Did students engage in self-evaluation?   [int_se]
    ❑ 1. Self-chart observations
    ❑ 2. Self-appraise performance
    ❑ 3. Self-administer primary reinforcers
    ❑ 4. Self-administer secondary reinforcers
    ❑ 5. Other steps _____
    ❑ 99. Not specified

E5. Did teacher monitor behavior?   [int_tchmon]
    ❑ 1. Yes
    ❑ 2. No_____(list)
    ❑ 99. Not specified

E6. What source did students use to compare SM data?   [int_tchcom]
    ❑ 1. Teacher
    ❑ 2. Self-goal
    ❑ 3. Prior performance
    ❑ 4. Other third party
    ❑ 5. Multiple sources: _____(list)
    ❑ 99. Not specified

E7. Were students provided training?   [int_trn]
    ❑ 1. Yes
    ❑ 2. No
    ❑ 99. Not specified

E8. Length of training   [int_lgth]
    ❑ Total number of minutes: _____

❑ 99. Not specified

E9. Teaching modalities used in training [int_mod]
  ❑ 1. Sequenced training (vertical alignment of skills)
  ❑ 2. Activities used in training (modeling, practice, rehersal)
  ❑ 3. Focused on SM skills (examples/nonexamples)
  ❑ 4. Explicit (lessons, manual)
  ❑ 99. Not specified

E10. Length of self-monitoring [int_lgth]
  ❑ # of weeks: _____
  ❑ 99. Not specified

E11. Focus of goal for behavior intervention [int_assess]
  ❑ 1. Increase display of positive behavior
  ❑ 2. Reduce display of challenging behavior
  ❑ 3. Increase positive and reduce challenging behaviors
  ❑ 4. Academic performance
  ❑ 5. Both behavior and academic outcomes
  ❑ 6. Other_____

E12. Manner of prompt used in the study? [int_prpt]
  ❑ 1. Internal prompt (self-prompt)
  ❑ 2. External prompt (third party verbal prompt)
  ❑ 3. Technology (assistive device; recording or digital)
  ❑ 4. Combination of internal and external prompt
  ❑ 99. Not Clear

E13. Type of recording device [int_redev]
  ❑ 1. Digital device only
  ❑ 2. Paper-and-pencil only
  ❑ 3. Assistive device and paper and pencil
  ❑ 4. Other _____
  ❑ 99. Not clear

E14. Behavioral Goal [int_goal]
  ❑ 1. Positively worded (presence, positive replacement behavior)
  ❑ 2. Negatively worded (absence, cease a challenging behavior)
  ❑ 3. Positively and negatively worded goals
  ❑ 4. Focus on off/on task behavior
  ❑ 99. Not clear

E15. Number of intervals on recording device_____ [int_interval]

E16. Number of recording device response options [int_resopt]
  ❑ 1. One
  ❑ 2. Tw0
  ❑ 3. Three
  ❑ 4. Four
  ❑ 5. Five
  ❑ 6. Six or more
  ❑ 99. Not clear

E17. Number of self-evaluation sessions [int_seses}
  ❑ 1. One
  ❑ 2. Tw0
  ❑ 3. Three
  ❑ 4. Four
  ❑ 5. Five
  ❑ 6. Six or more
  ❑ 99. Not clear

E18. What type of reinforcement used along with SM? [int_reinf]
  ❑ 1. None
  ❑ 2. Positive reinforcement

❑ 3. Negative reinforcement
❑ 4. Both positive and negative reinforcements
❑ 99. Not clear

E19. Was the intervention faded?   [int_fad]
❑ 1. Yes
❑ 2. No
❑ 99. Not specified

Section F—Intervention Fidelity

F1. Did the study report fidelity?   [fidel]
❑ 1. Yes
❑ 2. No
❑ 99. Not specified

F2. How was fidelity assessed?   [fidel_asses]
❑ 1. Researcher observations
❑ 2. Interviews of participants
❑ 3. Surveys of participants
❑ 4. Participant logs
❑ 5. Administrative records
❑ 6. Checklists
❑ 7. Other _____
❑ 99. Not specified

F3. Level of adherence to the tx: [fidel_ad]
❑ 1. Percent or Level _____
❑ 99. Not specified

F4. What type of fidelity measures were collected   [fidel_typ]
❑ 1. Structural (count of services delivered)
❑ 2. Procedural (quality of services delivered)
❑ 3. Both structural and procedural
❑ 99. Not specified