THE ROLE OF PERCEIVED SELF-EFFICACY IN CONTROLLING BEHAVIOR

By

Jennifer E. Samson

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

August, 2009

Nashville, Tennessee

Approved:

Joseph Wehby

Bethany Rittle-Johnson

Bahr Weiss

To Doug –

for everything

ACKNOWLEDGEMENTS

The data for this paper was collected by the Vanderbilt Behavior Research Center, in conjunction with the project "Reducing Severe Problem Behaviors in Schools." The project is supported by the Institute of Education Sciences, U.S. Department of Education through Grant H324P040013 to Vanderbilt University. I am supported by a predoctorial training grant provided by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B040110 to Vanderbilt University. The opinions expressed are mine and do not represent views of the U.S. Department of Education.

I am indebted to a huge number of individuals who made this paper possible. First, thanks to the students and staff at the Vanderbilt Behavior Research Center, as well as to the faculty, staff, and students at the schools that participated in the project. Thanks to Andrew Tomarken, Elizabeth Lemerise, Nancy Eisenberg, and the faculty of the Vanderbilt Department of Psychology and Human Development for taking the time to discuss components of this paper. Also thanks to my committee members – to Bahr Weiss and Bethany Rittle-Johnson for their careful readings and insightful comments, and to Joe Wehby for endless hours of discussion, guidance, and encouragement. Finally, thanks to friends and family, too numerous to name here, without whom my sanity would have been lost long ago.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The high prevalence of externalizing behavior problems such as aggression and delinquency, and the negative outcomes associated with these behaviors, have resulted in a large amount of research devoted to understanding the causes of these problems and developing methods for intervention. One factor that this research suggests might be important in the development of externalizing behavior problems is a sense of self-efficacy for control of emotionally-driven behaviors. The current study examined the relationship between self-reports of perceived self-efficacy for controlling emotionally-driven behavior and actual externalizing behavior problems among children with high levels of externalizing problems.

The following literature review is divided into three main sections. The first section contains a brief review of externalizing behavior problems, including a discussion of the prevalence of these behaviors and their associated negative outcomes. The second section focuses on perceived self-efficacy. Here, the relationship between perceived self-efficacy and achievement is reviewed. This review is followed by a description of existing literature relating behavior to perceived self-efficacy for inhibiting aggression, as well as an explanation of how perceived self-efficacy fits into a Social Information Processing (SIP) model (Crick & Dodge, 1994). In the third section, emotion regulation is introduced as an additional factor explaining externalizing behavior problems, both directly and as part of a SIP model. The combined literature suggests that beliefs about

self-efficacy for inhibiting aggression can be reconceptualized as beliefs about self-efficacy for controlling emotionally-driven behaviors. The section concludes with the suggestion that an explicit mention of emotion was missing from previous research, and a review of existing literature supporting the hypothesized relationship between perceived self-efficacy to control behavior and actual behavior.

Externalizing Behavior Problems

Externalizing, or "acting out," behavior problems (Kauffman & Landrum, 2009, p. 430) may include defiance, impulsivity, disruptiveness, aggression, and hyperactivity (Hinshaw, 1992), as well as rule-breaking behavior such as stealing and lying (Kauffman & Landrum). For toddlers, such behaviors are considered typical expressions of frustration or attempts at autonomy. The majority of children grow out of these behaviors, and they are considered problematic only to the extent to which they continue beyond toddlerhood or replace the development of age-appropriate social and communication skills (Campbell, Shaw, & Gilliom, 2000).

Some children, however, exhibit high levels of externalizing behavior throughout their childhood and into adolescence. Without intervention, children who display higher levels of externalizing behaviors than their peers are at risk for continued difficulties (Kazdin, 1995). In their most extreme forms, these types of behaviors may be indicators of disabilities such as ADD/ADHD, Conduct Disorder, and Emotional/Behavioral Disorder (American Psychiatric Association, 1994; "Individuals with Disabilities Education Act," 2004). Children with such disabilities often need special assistance in school. In 2004, almost 500,000 U.S. children required special education services for

Emotional Disturbance alone. These services ranged from extra support within a typical classroom to placement in a self-contained classroom or even in a separate school (*28th Annual Report to Congress,* 2009).

Even those children who display subclinical, but still higher than average, levels of externalizing behaviors are at risk. The difficulties these children face may include rejection by their peers, as well as loneliness (Asher & Paquette, 2003) and internalizing problems such as depression and withdrawal (e.g. Coie, Lochman, Terry, & Hyman, 1992). They are also at risk for continuing externalizing behavior problems (e.g. Hymel, Rubin, Rowden, & LeMare, 1990), including delinquency in adolescence (e.g. Kupersmidt & Coie, 1990) and criminality in adulthood (e.g. Huesmann, Eron, Lefkowitz, & Walder, 1984). Finally, children who exhibit high levels of externalizing behavior often experience academic difficulties, including school dropout (e.g. Ollendick, Weist, Borden, & Greene, 1992).

Externalizing behavior problems and negative outcomes can exasperate one another and, over time, can result in increasingly harmful outcomes. For instance, Dodge and colleagues (e.g. Coie, Dodge, Terry, & Wright, 1991; Dodge, 1983) found that children who displayed high levels of aggression were often rejected by their peers, but the causal pathways were unclear. Specifically, Dodge (1980) described a downward spiral in which children who experience rejection come to expect hostility from others and thus interpret others' actions as hostile. This interpretation causes the children to respond aggressively even to ambiguous (i.e., unclear if it was "on purpose" or not) social initiations. In turn, the child's peers are then more likely to label him/her as aggressive and act aggressively towards him/her in the future. This type of interaction is

but one example of how behavior problems and negative outcomes can worsen over time. Therefore, the goal of practitioners and researchers should be to interrupt this spiral as early as possible through intervention.

Although it has been found that the frequency, severity, and pervasiveness of behavior problems across contexts make it possible to identify children at risk for later problems as early as age 2 or 3 (Campbell et al., 2000), intervention for externalizing behavior problems, even when begun in preschool, has shown only limited success. In an early review of social-cognitive interventions, Urbain and Kendall (1980) found that results were mixed, especially in regard to how well the interventions affected actual behavior (as opposed to specific social cognitive skills).

Intervention results have improved only moderately since then. Recently, Wilson, Lipsey, and Derzon (2003) conducted a meta-analysis of school-based interventions for aggressive behavior. The meta-analysis revealed that interventions were generally successful. However, the vast majority of studies were conducted on a "demonstration" scale and those interventions that were scaled up to be incorporated into school curricula with less researcher support were much less effective. Overall, children who are highly aggressive in preschool stand a very good chance of struggling with aggression throughout their childhood and sometimes beyond. In fact, Kazdin (1995) reported that the correlation between aggression in preschool and at age 10 is higher than that for IQ and suggests that interventions after age 8-10 should focus more on management rather than remediation.

Perceived Self-Efficacy and Social Information Processing

Given the limited success of current interventions and the long-term impact of behavior problems, the challenge to the research community is twofold: 1) to continue the testing of interventions that can have a significant impact on problem behavior and 2) to extend descriptive research to help identify factors that influence problem behavior which may lead to the development of more effective interventions. The current study addresses the latter goal by focusing on one area that is lacking in our understanding of problem behavior – the role of beliefs about self-efficacy for inhibiting problem behavior.

Perceived self-efficacy can be defined as one's beliefs about his/her ability to succeed (Bandura, 1977). Bandura and colleagues (Bandura; Bandura, Pastorelli, Barbaranelli, & Caprara, 1999) suggest that perceived self-efficacy may predict, or even contribute to, certain actions. Specifically, people tend to report higher confidence in their ability when their ability is high and to choose actions where they feel confident in their ability to succeed.

The causal relationship between perceived self-efficacy and ability appears to be bidirectional. Bandura (1977) hypothesized that one major source of perceived self-efficacy is performance accomplishment, or the experience of success. Perceived self-efficacy has also been shown to predict success above and beyond previous ability (e.g. Lent, Brown, & Larkin, 1984; Lent, Brown, & Larkin, 1986; Meier, McCarthy, & Schmeck, 1984), suggesting that perceived self-efficacy has a relationship with achievement above and beyond being a report of reality. For instance, having high perceived self-efficacy for a particular action may cause one to choose that behavior often and thus become practiced at it, improving ability.

It has been reported that individual differences in success are correlated with individual differences in perceived self-efficacy in a wide variety of contexts (Schunk, 1989). Perceived self-efficacy has been studied extensively as a correlate of academic achievement, and to a lesser degree in relation to motor skills (diving, tennis, gymnastics, reaction time, muscle strength, and endurance), social competence (see Schunk, 1989 for a review), and health, including smoking cessation (DiClemmente, 1981). Perceived self-efficacy has been found in all of these domains to predict achievement.

Perceived self-efficacy predicted academic achievement for students from elementary school (Anjum, 2006; Bandura et al., 1999) to college (Lent et al., 1984; Meier et al., 1984) in writing (Meier et al., 1984), mathematics (Anjum; Norwich, 1987; Randhawa, Beamer, & Lundberg, 1993), and on more general cognitive tasks (Locke, Frederice, Lee, & Bobko, 1984). In the social domain, Coleman (2003) found a positive correlation between perceived social self-efficacy and attachment to peers among fifth and sixth graders. Connolly (1989) found that high schoolers' reports of perceived social self-efficacy were correlated with their reports of social acceptance and their parents' reports of social competence. In a longitudinal study, Bandura and colleagues (Bandura et al., 1999) found that early adolescents' perceived social self-efficacy predicted prosocial behavior. Finally, in a study with young adult women, Lee (1984) found that participants' self-efficacy for assertiveness in social situations predicted their performance during a role-playing task. The relationship between perceived self-efficacy and this wide variety of achievement outcomes suggests that perceived self-efficacy for inhibiting aggression might be an important predictor of success in social problem-solving.

The current literature review found two studies which report a relationship between problem behavior and perceived self-efficacy for inhibiting aggression. Perry, Perry, and Rasmussen (1986) used an interview based on vignettes portraying hypothetical provocations and found that children who had been rated by their peers as highly aggressive reported lower perceived self-efficacy to inhibit an aggressive response than their low-aggressive peers. Later, Gottheil and Dubow (2001) found, using an adaptation of the same measure, that perceived self-efficacy to inhibit aggression was lower among children classified as "bullies" (using peer and self-report measures) than among children not classified as "bullies."

Perry and colleagues' work is grounded in Social Information Processing (SIP) theory (Crick & Dodge, 1994), which suggests that behavior problems are related to deficits in social problem-solving. The presence of individual differences in social problem-solving proficiency has been well established in the cognitive-behavioral literature (e.g. Camodeca & Goossens, 2005; Dodge, 1980), and deficiencies have been linked to deficits in social competence, including externalizing behavior problems (e.g. Dodge & Price, 1994). For instance, misinterpreting the intent behind a social act and failing to generate multiple solutions to a problem have both been associated with increased aggression (VanOostrum & Horvath, 1997). Social Information Processing hypothesizes that perceived self-efficacy to enact a strategy is one factor in children's decisions regarding which strategy to employ in a social problem-solving situation (Crick & Dodge, 1994). The work of Perry et al. (1986) extended this theory by also examining perceived self-efficacy for *inhibiting* particular responses.

The Role of Emotion and Its Regulation

Perry et al. (1986) and Gottheil and Dubow (2001) suggest that a relationship exists between perceived self-efficacy for inhibiting aggressive behavior and actual behavior, but they do not attempt to explain *why* inhibiting aggressive behavior might be related to the ability to produce appropriate behaviors in a social situation. Since the publication of Perry et al.'s article, however, researchers have begun to examine the role of emotion and its regulation in successful processing of social situations.

Emotion regulation may be defined as the processes, both internal and external, which recognize and moderate the intensity and duration of emotional reactions (Thompson, 1994). The regulation of emotion-related behaviors (henceforth behavioral regulation) is often considered a part of emotion regulation and may be defined as the matching of emotional expression to environmental requirements (Walden & Smith, 1997). Behavioral regulation is nearly impossible to separate empirically from the regulation of emotion itself (Eisenberg, Guthrie, et al., 2000) and includes the masking of inappropriate expression and/or the display of appropriate expressions (Walden & Smith). Although the current paper will focus on conscious elements of emotion and behavioral regulation, it is important to remember that these processes may be, in part, subconscious (Walden & Smith), and that any intervention focusing on improving conscious emotion regulation may only be affecting part of the process.

The relationship between behavior problems and individual differences in regulation is well-established using a wide variety of methodology. Research has shown both concurrent and long-term predictive relationships between emotion and behavioral regulation and behavior problems (see Eisenberg, Fabes, Guthrie, & Reiser, 2000 for a

review). High levels of emotional expression in response to frustration, as demonstrated in lab tasks, has been shown to predict parent-reported behavior problems among children as young as 2 (Calkins, Gill, Johnson, & Smith, 1999). Eisenberg and colleagues (Eisenberg et al., 1993) found teacher and parent ratings of emotionality, negative affectivity, and acting out to be positively correlated among preschool children. Similar relationships have been demonstrated between regulation and externalizing behavior problems among first graders (Eisenberg et al., 2001) and fourth graders (McDowell, O'Neil, & Parke, 2000).

Researchers have also found long-term predictive relationships between regulation and behavior. For instance, using physiological signs (heart rate, vagal tone, skin conductance) and facial expression measures of emotionality, Cole and colleagues (Cole, Zahn Waxler, Fox, Usher, & Welsh, 1996) grouped preschoolers into "expressive," "inexpressive," and "modulated" groups. Halfway through first grade (2 years later), teachers of the expressive group reported that they displayed more externalizing behavior problems than the other children. Additionally, Eisenberg and colleagues have demonstrated a relationship between regulation and both teacher (Eisenberg, Guthrie, et al., 2000) and parent (Eisenberg et al., 1995; Eisenberg, Guthrie, et al., 2000) reports of later behavior problems.

The connection between regulation and processing of social situations has been described by Lemerise and Arsenio (2000), who suggested that emotional and behavioral regulation predict social competence because dysregulation interferes with competent processing. In this view, failure to successfully regulate emotion and behavior is seen as one factor leading to deficits in processing social problems. For example, unregulated

emotion may lead a child to misinterpret cues about the intent of a perceived provocation, to generate less prosocial strategies, or to fail to enact a chosen strategy competently. Lemerise and Arsenio emphasize that disruption by regulation difficulties may occur at any point of processing.

Research has supported emotional regulation's role in the processing of social information. For example, the degree to which a child is distressed by negative situations (Crick, Grotpeter, & Bigbee, 2002) and experimentally inducing a negative mood (Orobio de Castro, Slot, Bosch, Koops, & Veerman, 2003) have both been shown to predict a tendency to attribute hostility in ambiguous situations. In a second example, children who were normally highly aggressive generated and chose more socially acceptable responses to provocation when forced to slow down and acknowledge their own emotional responses (Orobio de Castro, Bosch, Veerman, & Koops, 2003).

Lemerise and Arsenio (2000) argue that emotion and its regulation make contributions to social problem solving which are in addition to cognitive processes described by Dodge and colleagues' earlier model (Crick & Dodge, 1994). Because emotion regulation affects every step of the problem-solving process, associated cognitions, including perceived self-efficacy to regulate emotion and/or behavior, may also affect every step. In other words, a child who does not believe he/she can regulate an initial, emotional response and proceed with more competent processing of the social problem at hand may not even try to do so or may resort to established patterns of maladaptive behavior.

However, the role of perceived self-efficacy to regulate emotion and emotionally-driven behavior has been largely ignored. The vignettes used in Perry et al.'s (1986) and

Gottheil and Dubow's (2001) studies do not reference the negative emotional reactions that might cause one to respond aggressively to provocation or how inhibiting an aggressive response may involve first regulating negative emotions. Because dysregulation affects competent social processing, it may very well be that children reason differently about their perceived self-efficacy to inhibit aggression when the emotion driving the aggression is made explicit than when it is not. For example, it has been shown that children perceive more hostility in provocations which they experience as more distressing (Crick et al., 2002), but it is not known whether distress causes hostile attribution or vice versa. Thus, it is possible that knowing they are (hypothetically) upset by the provocation in a vignette may cause a participant to report more hostile attribution than if their emotional reaction is left unstated.

While the relationship between perceived self-efficacy to inhibit emotional expression and behavior has not been studied directly, previous research does suggest that a relationship might exist. As discussed above, perceived self-efficacy predicts success in a wide variety of other domains (e.g. as reviewed in Schunk, 1989). Also, perceived self-efficacy to regulate emotional reactions appears to increase at about the same time most children begin to succeed at regulating these behaviors. Specifically, in an earlier study examining age and gender patterns in children's perceived self-efficacy to control behavior and emotion (Samson & Giles, manuscript in preparation), it was found that reasoning about one's ability to control emotion and behavior differs between preschool and elementary school children. While children of all ages believed that it would be easier to change a behavior than the emotion underlying it, older children and adults reported more control over their behavior than younger children. These changes in

perceived self-efficacy coincide with the age at which children begin to understand hiding emotions and to consistently use display rules – suppressing or replacing an emotional display in order to present a more socially acceptable display (Saarni, 1984) – to regulate their behavior (see Harris, 1989 for a review). In addition, these age-related changes in perceived self-efficacy match age-related changes in what is considered "typical" levels of externalizing behaviors (e.g. behaviors such as tantrums that are considered problematic in an 8-year-old but are typical for a 3-year-old).

Thus, the current study attempts to extend the self-efficacy literature to include perceived self-efficacy to regulate emotionally-driven behaviors and to extend the behavior problem literature by exploring the role of perceived self-efficacy to regulate emotionally-driven behaviors in predicting behavior problems. Specifically, the goal of the current study is to relate individual differences in externalizing behaviors to individual differences in perceived self-efficacy to control emotionally-driven behavior, where the emotional origin of the behavior is explicitly stated. The current study also expands on Perry et al. (1986) by examining less extreme forms of externalizing behaviors (Perry et al. asked only about perceived self-efficacy to inhibit aggression) and by using teacher and observer reports of behavior rather than peer reports (see "Measures" section below for a discussion of the advantages of using multiple methods to measure behavior). Finally, because the end goal for this line of research is intervention, the current study is focused on children for whom pathways to intervention for externalizing behavior problems are sorely needed – those receiving special education services for emotional and behavioral disorders or those at risk for needing these services.

Measures

One limitation of previous research relating perceived self-efficacy for inhibiting aggressive behavior to actual behavior is the reliance on self or peer reports to measure behavior. In a review of best practices in assessing children's social skills for research and education purposes, Merrell (2001) concludes that a combination of observation and rating scales should be the "first-line" approach for evaluating children's social skills. Thus, the use of observation and teacher ratings to quantify behavior in the current study is an important addition to the literature.

Rating scales offer a quick, efficient snapshot of a child's behavior and may be useful as a first step in identifying children for whom intervention (or at least further evaluation) is necessary. The Child Behavior Checklist – Teacher Report Form (CBCL-TRF, Achenbach & Rescorla, 2001) used in the current study offers a relatively quick (about 15 minutes per child) picture of the general tendencies of participants' behavior. The CBCL-TRF demonstrates good reliability and validity (Achenbach & Rescorla), and is, in fact, often used to validate new measures (e.g. Caldarella, Young, Richardson, Young, & Young, 2008). However, it is not without flaws. Specifically, the CBCL-TRF is based on one teacher's view of each participant's behavior and is thus subject to bias. Because each participant's primary teacher (i.e. different teachers across participants) completed the rating scale (see details in "Method" section below), these biases may or may not be the same for each participant.

On the other hand, direct observation may offer a more objective view of behavior. Because observers have undergone training and have had interrater reliability assessed and maintained (see details in "Methods" section below), and because their

knowledge of the participant outside the observation is limited (hence the term, "objective" ratings), observation is less subject to individual biases. However, observation is more time-consuming and expensive and may capture a different view of behavior than a rating scale because even a valid, reliable coding system only rates small slices of behavior. Specifically, an observation system may miss low-frequency, highly important behavior (e.g. a tantrum once a week is disruptive to learning but may be missed by observation if it occurs outside the time behavior is sampled). The current observation platform, *Multiple Option Observation System for Experimental Studies* (MOOSES; Tapp, Wehby, & Ellis, 1995), has been used previously to describe the behavior of students with special needs (e.g. Carter, Sisco, Brown, Brickham, & Al-Khabbaz, 2008; Shores & Wehby, 1999). Given the strength of observation for objective coding of small samples of behavior and the strength of rating scales for capturing more global behavior patterns, using both measures in the current study should allow a more complete picture of participants' externalizing behavior problems than was presented in previous research.

The final instrument used in the current study is an experimenter-designed measure of children's perceived self-efficacy to control their behavior. The *perceived self-efficacy to control behavior* measure included in this analysis was originally developed for another project to describe age and gender patterns in perceived self-efficacy to control emotion and emotionally-driven behaviors (Samson & Giles, manuscript in preparation). The vignettes used in the measure (see "Method" section for details) are adapted from an established measure of children's understanding of hidden emotion (i.e. the possibility of experiencing one emotion internally but expressing a

different emotion; Harris & Gross, 1988), with only minor changes (e.g., third person to second person).

The measure shows high face validity, as its main question asks specifically about the construct of interest – perceived ability to control behavior. The wording of the perceived self-efficacy question ("Could you stop yourself from ____?") is identical to that used in previous studies, in which children's understanding of the question was never an issue. In addition, previous work with this measure suggests construct validity because results have been consistent with established theory about when children develop an understanding of hidden emotion and, more generally, theory of mind – that is, a difference between preschoolers' and elementary schoolers' reasoning. Although reliability has not been established for the *perceived self-efficacy* measure, its use in the current exploratory study seems justified, and the current study may offer insights into ways to improve the measure.

CHAPTER II

PURPOSE

The current study aims to build on the growing body of research to explain behavior problems by examining the role of perceived self-efficacy to regulate emotional outbursts that take the form of externalizing behavior problems. The current study also extends this work to focus on children for whom behavior problems are more severe than average – those with or at risk for emotional or behavior disorders (EBD). Finally, to get a more complete picture of participants' behavior, the current study uses both teacher and observer reports. In short, the primary goal of the current study is to begin to assess the nature of the relationship between children's perceived self-efficacy to control emotionally-driven behavior and their actual behavior; specifically, the extent to which self-efficacy beliefs regarding one's ability to control one's behavior predict observed disruptive behavior and/or teacher ratings of problem behaviors. Based on previous research (Gottheil & Dubow, 2001; Perry et al., 1986) and a theoretical relationship between perceived self-efficacy and ability in a variety of domains (Bandura, 1977; Schunk, 1989), it is hypothesized that higher perceived self-efficacy to control behavior will predict lower levels of problem behaviors, as measured by teacher and observer reports.

CHAPTER III

METHOD

The present data was collected as part of a larger, federally funded project, the Vanderbilt Behavior Research Center, whose purpose was to experimentally test a package of evidence-based classroom practices designed to improve the reading achievement and behavior of students with and at risk for Emotional or Behavior Disorder (EBD). The project included children in both regular education classrooms and self-contained classrooms – an environment especially designed for children with behavior problems severe enough to prevent their success in a typical classroom. These self-contained classrooms feature a lower student to teacher ratio and specially trained teachers and offer more individualized academic and behavior management programs than can be delivered in a typical classroom.

Participants

Recruitment and Screening

Following district and principal approval of the project, teachers of general and special education students were informed of the purpose of the study and asked to consider participating. All teachers of elementary self-contained classrooms, and those general education teachers in the same schools that expressed interest, were contacted to obtain consent. A total of 113 teachers agreed to participate.

To identify the children at highest risk for behavior problems in general education classrooms, the multi-gated *Systematic Screening of Behavior Disorders* (SSBD, Walker & Severson, 1990) was used. Teachers were first given a description of a student with externalizing behavior problems and asked to identify the three students in their classroom who best fit the description. The research team provided letters to be sent home asking for parent consent to screen students in participating classrooms for eligibility to participate in the project. A targeted goal of 80% returned positive consents was met in 67% of the classrooms for a total of 945 children screened for eligibility using the SSBD.

Once parent consent for screening was received, the teachers completed two sections of the screening measure for each identified child. The first section identified Critical Events and the second identified the frequency of adaptive and maladaptive behaviors. Those students who were reported to have either (a) five or more Critical Events or (b) at least one Critical Event, an Adaptive Behavior score less than 30, and a Maladaptive Behavior score of at least 35 were considered eligible to be included in the project. These predetermined cut scores indicate the presence of clinical levels of behavioral issues (Walker & Severson, 1990).

In self-contained classrooms, all students were automatically eligible for the study by virtue of their placement in special education classrooms specifically designed for children with behavioral issues. Parent consent to participate in the project was obtained for general education students who met these criteria and for all special education students.

Current Sample

The current data were collected at one time point from one of the study sites. Two children from the site were excluded from this analysis because of incomplete data, leaving a final sample of 129 children (age 5y8m – 11y 9m, 20.2% female) from 49 classrooms. Seventy-three (56.6%) of the children had been placed in self-contained classrooms specifically designed for children with behavioral issues. Twenty-nine students had special education labels but were being served in a regular education classroom. Although all of these children demonstrated high levels of behavior problems, the special education labels under which they were being served varied (see Tables 1a-b). The remaining participants were considered at-risk for needing special education services for behavior problems.

Table 1a: Primary Disability Category by Placement

	ADD/ ADHD	autism	dev't delay	EBD	health	learning disabled	mentally retarded	speech/lang impairment	n/a or unknown
reg ed	3	1	7	6	1	5	1	5	27
self-contained	13	0	10	22	8	7	2	10	1

Table 1b: Secondary Disability Category by Placement

	ADD/ ADHD	EBD	health	learning disabled	speech/lang impairment	n/a or unknown
reg ed	0	1	1	1	1	52
self-contained	1	0	1	0	2	69

Procedure

The data used in the present analysis were collected in the spring of the school year and included (a) an experimenter-designed measure of student's self-reported perceived self-efficacy, (b) the Child Behavior Checklist – Teacher Report Form (Achenbach & Rescorla, 2001), and (c) observation of the students' behavior using the MOOSES (Tapp et al., 1995) observation platform.

Perceived Self-Efficacy to Control Behavior

The *perceived self-efficacy* measure was presented in individual interviews by trained research assistants. Children were pulled from their class, at a time convenient for the teacher, for about 20 minutes (the *perceived self-efficacy* measure took 5-10 minutes; an additional emotion recognition measure, reported elsewhere, was conducted during the same interview). These interviews were conducted in an empty classroom, with three to four experimenter-child pairs working simultaneously. These pairs were seated at different tables to minimize distractions.

The *perceived self-efficacy* measure consisted of two short vignettes (see Appendix A) which the experimenter read out loud to the participant. In each vignette, participants imagined themselves in a situation where they experienced a specific emotion (happy or upset) and had an opportunity to control an associated behavior. They were always presented with the reason for the emotion as well as for needing to control the behavior associated with that emotion. For instance, in one vignette, the participant imagines he/she won a game and feels happy enough to cheer. However, the participant is reminded that if he/she cheers, his/her friend will be not want to play anymore.

Following presentation of each vignette, participants answered a *perceived self-efficacy to control behavior* question. This forced-choice question examined the participant's perception of his/her ability to control his/her actions (e.g., "Could you stop yourself from cheering?"). If the participant answered "yes," perception of difficulty was assessed with the follow-up question, "Would it be easy or hard to stop _____?" "No" responses were given a score of 0, "hard to stop" responses were given a score of 1, and "easy to stop" responses were given a score of 2, so that higher scores indicated more perceived control. *Perceived self-efficacy to control behavior* scores were computed for each participant by summing across the two vignettes, so each score had a minimum value of 0 and a maximum of 4. (Note that the participants also answered a *perceived self-efficacy to control emotion* question following each vignette, which was dropped from this analysis because it was outside the scope of the current research questions.)

Teacher Report

The Child Behavior Checklist – Teacher Report Form (CBCL-TRF; Achenbach & Rescorla, 2001) was completed by each participant's primary teacher. The CBCL-TRF is a 112 item questionnaire. Each item asks about the frequency with which the teacher observes a particular behavior (e.g. cries easily, destroys property, swears), with three response options: not true, somewhat true, and very true. The instructions on the CBCL-TRF report form do not specify to the teacher *where* the behaviors might be observed. It is assumed that it measures the teacher's general impression in all school settings, but is weighted heavily by the teacher's experience with the student in classroom interactions. The scale yields a profile of the child's behavior, including broadband internalizing and

externalizing behavior problems, which are then further broken into subscales. The current analysis uses both the *Externalizing* scale and the *Internalizing* scale, along with *Externalizing* subscales *Aggression* and *Rule-Breaking*. The *Social Skills* scale (which is separate from both broadband scales) was also used for this analysis.

Observation

Trained observers conducted classroom observations using the *Multiple Option Observational System for Experimental Studies* (Tapp et al., 1995). MOOSES is a computer-based observation system for simultaneous collection of discrete events and durational measures as they occur in real time. Behavioral variables associated with (a) students, (b) teachers, and (c) instructional setting were collected using MOOSES. Research assistants conducted four 15-minute observations per student, with data pooled across these sessions.

Most of the observation data focused on the interaction between teacher and student, which is irrelevant to the current analysis. Thus, only the student variable *disruptive behavior* was used here. The codes *negative talk*, defined as "statements or vocalizations made with the intent to provoke, annoy, pester, mock, whine, complain, tattle, or make fun of another" and *aggressive behaviors*, defined as "deliberate physical contact that is potentially harmful to self, others or property [or] posturing or a gesture that is intended to provoke another" were combined to create the variable *disruptive behaviors*. This variable reported the count of disruptive behaviors within the 60 minute observation period. (See Appendix B for complete coding definitions.) In cases where observation did not occur for exactly 60 minutes, the rate per minute was calculated and

then multiplied by 60 to project how many instances of disruptive behavior might have been observed in exactly 60 minutes. These projections were then rounded to the nearest whole number for count data analyses that required integers.

Before collecting the observation data, research assistants underwent rigorous training to ensure reliability. Observers-in-training first reviewed the descriptions of behavioral codes. Next, observers-in-training practiced coding 15-minute video clips that had been previously scored by a master coder. The observer-in-training was required to reach reliability of .80 with the master coder on three consecutive videos before moving on to live coding practice. For live practice, the observer-in-training was paired with a trained coder for 15-minute coding sessions in a nonparticipating classroom. Once they reached reliability of .80 with a trained coder on three consecutive live practices, the new observer was deemed reliable.

Ongoing reliability was assessed by pairing research assistants to simultaneously complete observations approximately once a week (for 20% of observation sessions). Interrater agreement was calculated using the traditional percent agreement formula [agreements ÷ (agreements + disagreements)] * 100. Mean agreement for disruptive behavior was 98.3%. Percent agreement was used because the observers were coding instances of specific types of behavior within a time period rather than categorizing behaviors as would be required to compute Cohen's kappa.

CHAPTER IV

RESULTS

The results will be presented in two sections. First, descriptive statistics will be presented for the teacher ratings, observation measures, and perceived self-efficacy. Then, *perceived self-efficacy to control behavior* scores will be entered into regression analyses to predict 1) observed behavior and 2) teacher reported behavior. Previous research suggests that age (Kauffman & Landrum, 2009), gender (Skiba, Poloni-Staudinger, Gallini, Simmons, & Feggins-Azziz, 2006), and placement in a self-contained versus a regular education classroom (Kauffman & Landrum) may predict externalizing behavior problems. Thus, these covariates are included in all analyses.

Descriptive Statistics

Child Behavior Checklist – Teacher Report Form

See Table 2 for descriptive statistics from the individual CBCL-TRF scales. The CBCL-TRF yields both raw and normed t-scores for each scale. Because the t-scores are standardized on a typical population, they may be inappropriate for this high-risk sample. However, they are reported here in order to allow the reader to orient this sample within a wider population. These t-scores can be categorized as typical, borderline, or clinical based on cutoffs suggested by normative data. The current sample, on average, scored in the borderline to low clinical range on the scales included in this analysis. (This was not

surprising given that they were selected for the project based on high levels of behavior problems.)

Table 2: Descriptive Statistics for the CBCL-TRF

Scale	Mean	Standard deviation	Borderline cutoff	Clinical cutoff
social	66.23	7.73	65	70
internal	62.78	9.64	60	63
external	68.09	9.54	60	63
external subscores				
aggressive	69.41	11.75	65	70
rule-breaking	65.36	8.95	65	70

All subsequent analyses use the raw scores for *externalizing behavior problems* (and subscales *aggression* and *rule-breaking*), *internalizing behavior problems*, and *social skills*. In each of these scales, the distribution of raw scores showed a skew towards lower scores. Thus, for all analyses, the scores were transformed by taking their square root to normalize the distribution.

Separate ANOVAs to predict *internalizing raw scores*, *social raw scores*, *externalizing raw scores*, and *aggressive raw scores* from age, placement, and gender were all nonsignificant (internal: $F_{7, 121} = .72$, ns; social $F_{7, 121} = .42$, ns; external $F_{7, 121} =$

1.23, ns; aggressive $F_{7, 121} = .88$) indicating no group differences in these variables. However, an ANOVA on *rule-breaking* scores (model $F_{7,121} = 2.35$, $p < .05$; see Table 3) indicated an interaction of age and gender ($F_{1, 121} = 7.63$, $p < .01$), such that, among younger children, boys were rated as exhibiting higher levels of rule-breaking behavior than girls, but after age 8 this pattern reversed (see Figure 1). There was no effect for placement in a self-contained versus a regular education classroom.

Table 3: ANOVA on Rule-Breaking Scores

Source	DF	Type III SS	MS	F
age	1	5.36	5.36	6.27*
self-contained	1	.00	.00	0
female	1	6.27	6.27	7.33*
age*self-contained	1	.13	.13	.15
age*female	1	6.52	6.52	7.63*
self-contained*female	1	.42	.42	.49
age*self-contained*female	1	.40	.40	.47
within	121	103.39	.85	

* $p < .05$

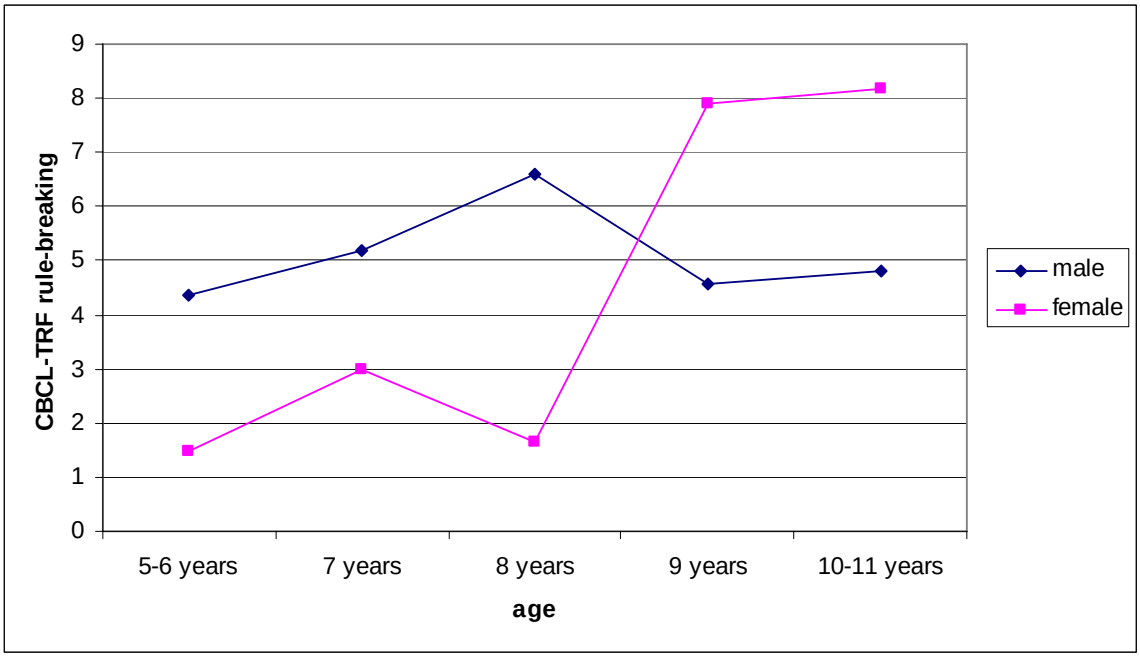

Figure 1: Male and Female Rule-Breaking Scores by Age

Observed Disruptive Behavior

The rate of observed *disruptive behaviors* per hour in the current data set ranged from 0 to 29 ($M = 2.18$, $s = 4.62$). The distribution was extremely skewed, with about 50% of observations equaling 0. However, because there is a large practical difference between children acting out once an hour and acting out 29 times an hour, the decision was made to maintain the original variability rather than converting the data to a binomial variable such as "observed disruptive behavior yes/no." Thus, an inflated-zero negative binomial regression was undertaken to examine the effects of age, gender, and placement on the observed count of disruptive behaviors.

An inflated zero negative binomial regression combines logistic modeling to analyze the zero/non-zero aspects of the data and negative binomial modeling to analyze the count data from one through 29 ("Stata annotated output"). Results (interpreted using

"Stata annotated output") indicated that the zero-inflated negative binomial model fits significantly better than a typical negative binomial model (Vuong $z = 1.85$, $p < .05$) or a Poisson model (lnalpha $z = 3.45$, $p < .01$), but that none of the independent variables (age, gender, and placement) significantly predicted *disruptive behavior* (LR $X^2 = 3.56$, ns). It is possible that the analysis lacked sufficient power to detect an effect – experts caution against using a zero-inflated model with "small" samples, but do not clarify what constitutes a "small" sample ("Stata data analysis examples").

Perceived Self-Efficacy to Control Behavior

Perceived self-efficacy to control behavior scores ranged from 0 to 4, with a mean of 2.95 (sd = 1.12). Because *perceived self-efficacy* was calculated as the sum of two items scored 0-2 (i.e., had limited possible values), a categorical distribution emerged. The distribution for *perceived self-efficacy to control behavior* displayed an increasing number of responses as scores increased (i.e., very skewed towards higher scores).

Thus, *perceived self-efficacy to control behavior* was treated in all analyses as count data, using negative binomial regression models that examine the distribution of scores around each point on the continuum of scores rather than the distribution of scores across the continuum. Additionally, the use of a negative binomial model allows the variance at each score to be larger than its mean, thus correcting for overdispersion as would be present in a Poisson model (Agresti, 2007).

To test for the possibility that children with low verbal skills may not have understood the *perceived self-efficacy* measure, exploratory analyses were undertaken to determine if those children with autism, mental retardation, or speech/language delay differed from the rest of the sample. Formal analysis was not possible due to low

numbers of children with autism and mental retardation (n = 4), so visual inspection confirmed that these children were not outliers. However, a t-test revealed that the mean *perceived self-efficacy to control behavior* score was lower among children with speech/language delay than in those without a delay (M = 2.28 versus 3.06 out of 4, $p < .05$). These children matched the rest of the sample in that they displayed above-average levels of behavior problems, and it was not clear that this difference was due to misunderstanding the measure. As a result, the decision was made to retain these children in the analysis but control for speech/language delay as a primary or secondary disability in subsequent analyses.

A negative binomial regression was performed (see Table 4) to examine the effects of age, gender, placement, and speech/language delay on *perceived self-efficacy to control behavior*. Results indicated that the effect for speech/language delay held even after controlling for age, gender, and placement (B = .23, $p < .05$), such that children with speech/language delay reported lower *perceived self-efficacy to control behavior* than those without speech/language delay. Additionally, there was a marginal effect for age (B = .06, $p < .10$) such that older children reported higher levels of *perceived self-efficacy to control behavior*.

Table 4: Negative Binomial Regression to Predict *Perceived Self-Efficacy to Control Behavior*

Variable	Estimate+	Chi-square
intercept	.23	.35
age	.06	2.95*
self-contained	-.10	2.53
female	-.05	.33
speech/language	-.24	5.51**

*p < .10 **p < .05

+Note that estimates are not directly interpretable as effects on scores, but direction and significance are read as if this were linear regression.

Zero-Order Correlations

Pearson correlations were calculated between all variables (see Table 5). As might be expected, the five Child Behavior Checklist (CBCL-TRF) scores were all moderately to highly correlated with each other, and the subscales *aggressive* and *rule-breaking* were very highly correlated with the general scale *externalizing*. Observed *disruptive* behavior was not correlated with *perceived self-efficacy to control behavior*. In addition, the *social*, *aggressive*, and *externalizing* scales were moderately correlated with the observation variable *disruptive* and the *externalizing* and *rule-breaking* subscales were correlated with *perceived self-efficacy to control behavior*. However, note that the

correlation between perceived self-efficacy and the CBCL-TRF scores is positive – opposite the hypothesized relationship.

Table 5: Correlations

	self-efficacy	observed disruptive	CBCL external	CBCL rule-break	CBCL aggressive	CBCL internal	CBCL social	age	female	self-contained
self-efficacy	-	.00	.22*	.31*	.17	-.02	.10	.23*	.00	-.13
disruptive		-	.30*	.14	.33*	.05	.19*	-.14	-.10	.15
external			-	.84*	.98*	.31*	.53*	.04	.04	-.11
rule-breaking				-	.71*	.26*	.42*	.08	.08	-.15
aggressive					-	.31*	.53*	.01	.03	-.08
internal						-	.62*	.13	.08	-.01
social							-	.04	.07	-.03
age								-	.22*	.13
female									-	-.03
self-contained										-

*p < .05

Predicting Observed Disruptive Behavior

To test the hypothesis that *perceived self-efficacy to control behavior* would predict observed and teacher-reported behavior, a series of regression analyses were undertaken. The first analysis attempted to predict observer-reported *disruptive* behaviors from student-reported *perceived self-efficacy to control behavior* along with covariates age, self-contained, gender, and speech/language delay. An inflated-zero negative binomial regression was used (again due to the extremely skewed distribution of *disruptive*). Results (interpreted using "Stata annotated output") indicated that the zero-inflated negative binomial model fit significantly better than a typical negative binomial model (Vuong $z = 2.04$, $p < .05$) or a Poisson model (lnalpha $z = 3.33$, $p < .01$), but that none of the independent variables significantly predicted *disruptive behavior* (LR $X^2 = 4.54$, ns). As discussed above, it is unclear if this analysis was underpowered. However, given the lack of zero-order correlation between observed *disruptive* behavior and *perceived self-efficacy* reported above, it is more likely that no relationship exists than that this analysis was underpowered.

Predicting CBCL-TRF

The second regression analysis attempted to predict teacher-reported behavior from *perceived self-efficacy to control behavior* and the covariates age, gender, self-contained, and speech/language delay. Models attempting to predict *internalizing* and *social* were nonsignificant (internal $F_{5, 123} = 1.23$, ns; social $F_{5, 123} = 1.49$, ns). A model

predicting *externalizing* behavior was also nonsignificant ($F_{5, 123} = 1.66$, ns), as was a model predicting the *aggression* subscale score ($F_{5, 123} = 1.14$, ns).

On the other hand, a model predicting the *rule-breaking* subscale was significant (see Table 6, model $F_{5, 123} = 3.22$, $p < .01$). Note that this model is significant even if a Bonferroni correction is performed to guard against Type I error due to multiple analyses ($p < .05 \div 5$ analyses $= p < .01$, Maxwell & Delaney, 2004). Contrary to hypothesis, but in keeping with the results from the zero-order correlation analysis reported above, higher *perceived self-efficacy to control behavior* emerged as a unique predictor of a higher *rule-breaking* score ($B = .24$, $p < .01$), even after controlling for age, gender, placement, and speech/language delay. Additionally, being in a self-contained classroom marginally predicted a lower *rule-breaking* score ($B = -.30$, $p < .10$). Note that this finding is not unusual in this special education literature; see the discussion of possible teacher biases in "Discussion" section.

Table 6: Regress CBCL-TRF Rule-Breaking Scores on Perceived Self-Efficacy

	Estimate+
intercept	1.64**
control behavior	.24**
age	.01
female	.09
self-contained	-.30*
speech/language	.05
Model F	3.22**
R-squared	.1157

** p < .01 *p < .10

+Recall that this analysis used the square root of *rule-breaking* scores, so these estimates are not directly interpretable as actual effects on scores. However, direction and significance can be read as usual.

Post-Hoc Exploratory Analysis

Because previous research found a relationship between aggression and perceived self-efficacy for inhibiting aggression, it was also unexpected that the *aggressive* subscale was not related to *perceived self-efficacy*. It was hypothesized that the lack of relationship between *perceived self-efficacy to control behavior* and the *aggressive* subscale resulted from the *aggressive* subscale's inclusion of both emotionally-driven

reactive behaviors (e.g. where the aggression is a response to provocation) and proactive behaviors (e.g. where the aggression is an attempt to gain materials). However, an exploratory factor analysis failed to isolate the "emotion-driven" behaviors within the scale, possibly because most items are relatively ambiguous as to their cause. In addition, an "emotional aggression" subscale created with the items theoretically attributable to failure to regulate emotion was unrelated to *perceived self-efficacy to control behavior*. These analyses went beyond the scope of the scores generated by the CBCL-TRF, and no conclusions should be drawn unless they are replicated with an instrument designed to separate reactive from proactive aggression.

CHAPTER V

DISCUSSION

Summary and Explanation of Findings

The goal of the current study was to examine the relationship between children's perceived self-efficacy for behavioral regulation and their actual behavior, as reported by both teachers and objective observers. Previous work (Gottheil & Dubow, 2001; Perry et al., 1986) found that children's higher perceived self-efficacy to inhibit an aggressive response when faced with a provocation predicted their decreased reputation as either "aggressive" (Perry et al.) or a "bully" (Gottheil & Dubow). The current study attempted to build on this literature by making explicit to participants that inhibiting aggression is often equivalent to inhibiting an emotional response, and by asking about inhibiting less extreme forms of externalizing behavior (screaming and cheering at inappropriate times) rather than about inhibiting aggression itself. Also, the current study attempted to replicate this relationship among children with higher-than-average levels of behavior problems, as opposed to the typical population examined in previous work. It was hypothesized that, as in previous studies, higher perceived self-efficacy to control behavior would predict lower rates of teacher and observer-reported externalizing behavior problems.

The current findings confirmed that there is a relationship between perceived self-efficacy to control behavior and teacher-reported behavior problems. However, the

direction of this relationship is the opposite of that found in prior research. Specifically the current study found that higher levels of perceived self-efficacy for controlling behavior predicted higher levels of problem behaviors. Additionally, it was found that *perceived self-efficacy to control behavior* was related to teacher-reported *rule-breaking* behavior, but not to *aggressive* behavior or to observer-reported *disruptive* behavior. These results suggest that the relationship is not as simple as was originally hypothesized.

Direction of the Relationship between Perceived Self-Efficacy and Rule-Breaking

The current results suggested a direct relationship between perceived self-efficacy for behavioral control and teacher-reported rule-breaking behavior – a finding opposite that of the hypothesized relationship. The probability of this finding being due to statistical chance is low even by conservative standards. Why was this relationship positive in nature? Several explanations exist.

One possibility for the direction of the current findings is that the children in this sample over-reported their ability to control emotional expression. Perhaps these children have unrealistic or inaccurate perceptions of their abilities. Work in other domains has supported the idea that the lowest achievers might report the highest perceived self-efficacy. For instance, research has found a similar disconnect in reading and writing where the children who were the lowest in ability had the least realistic idea of how well they were doing and thus reported the highest perceived self-efficacy (Paris & Oka, 1986; Stipek, 1993). Schunk (1989) suggested that children using incorrect algorithms (and therefore getting incorrect answers) continue to incorrectly solve math problems and report high perceived self-efficacy, presumably because their strategies

"work" every time and the students were unaware that their answers were incorrect. It has also been shown that children who are aggressive overestimate their peer status (Hymel, Bowker, & Woody, 1993). Perhaps these children do not realize their strategies are "incorrect" and therefore continue to use them and to report high perceived self-efficacy.

In the current sample, students' beliefs about their abilities might also be influenced by their environment. In a classroom context, there are typically behavior parameters set forth by a teacher, and rewards and consequences in place to support appropriate behavior. These supports are especially likely to be present for the children in the current sample, who are known to exhibit high levels of behavior problems. With these extra supports in place, the children might feel that they can control their behavior, and therefore they might perceive more control than they actually demonstrate. In particular, because the children in the current sample exhibit higher levels of behavior problems than average, they have likely spent a large portion of their childhood being reprimanded for their behavior and adults have suggested multiple strategies for controlling their behavior. Because of these prior experiences, these children can list multiple strategies (e.g. take a deep breath, hit a pillow instead of your sister") that are believed to control behavior, when in actuality, these students are not effectively using those strategies. The relationship between their perceived self-efficacy and their behavior would then be positive because they can list more strategies than their typical peers, but exhibit more behavior problems because being able to list strategies is not effective unless they can *use* the strategies.

Finally, it is possible that the externalizing behaviors reported in the TRF actually are under the children's control, and hence the more control children had over the behaviors, the higher the level of the externalizing behaviors. Functional behavior assessment (FBA), a technique used to determine appropriate intervention for children with behavior problems, is based on the assumption that some children act out because their behavior is achieving something they consider rewarding (e.g. as reviewed in Fox & Davis, 2005). For instance, a child who feels ignored when he/she raises his/her hand to get attention may resort to calling out in class – a behavior that is effective in accessing attention but is labeled by the teacher as maladaptive.

Relationship between Perceived Self-Efficacy and Rule-Breaking but not Aggression

Regarding the lack of a relationship between perceived self-efficacy and the aggression subscale (which one might believe to be where most emotionally-driven, dysregulated behaviors might be reported), this is most likely an artifact of the measures used. First, the perceived self-efficacy scale is based on inappropriate expressions of emotion, such as yelling out, rather than aggressive behaviors. In this way, it is more closely aligned with the rule-breaking scale than the aggression scale (although the aggression scale does include some items such as yelling). In addition, the aggression subscale asks about a wide variety of behaviors (Achenbach & Rescorla, 2001), only about half of which might be considered the result of unregulated emotion (see description in "Results" section of exploratory analyses to separate "emotional" behaviors). It is believed that the misalignment of these measures contributed to the lack

of an empirical relationship between perceived self-efficacy and teacher-reported aggression.

Teacher-Reported versus Observer-Reported Behavior

Finally, the current findings indicated that, in spite of a relationship between perceived self-efficacy to control behavior and teacher-reported behavior, and between teacher-reported behavior and observed behaviors, there was no relationship between perceived self-efficacy and observed behaviors. This finding also may be a result of the methods used to measure behavior, specifically the different types of behaviors that may have been captured by the teacher report versus independent observers.

These two measures reported on different behaviors and in different contexts. The CBCL-TRF asks about teachers' experience with a wide variety of externalizing behaviors over several months in a variety of school settings (Achenbach & Rescorla, 2001), while the observation measure samples only specific instances of certain disruptive behaviors, occurring only during instruction, in a relatively short observation period. It is believed that the perceived self-efficacy measure predicted the CBCL-TRF scores but not the observation reports because the perceived self-efficacy and CBCL-TRF measures were both tapping more general behavior/cognitions while the observation reports tapped more specific behaviors. Additionally, if the observed relationship between the CBCL-TRF rule-breaking scale and perceived self-efficacy is correct, then the observation report would not have detected an effect because only aggressive behavior and one type of rule-breaking (yelling out in class) was coded.

Limitations

Given the exploratory nature of this study, several limitations should be acknowledged. Measurement issues, a limited sample, and the correlational nature of the study limit the conclusions which can be drawn.

Measurement

Although the current study attempted to improve on existing research by including both the CBCL-TRF and observations of behavior, it is important to remember that neither the teacher report nor the observer report is an entirely accurate representation of behavior. One example of possible bias in the CBCL-TRF is the finding that teachers reported higher levels of externalizing behaviors among females, even though it is well known that externalizing behavior problems are usually more severe in males (e.g. Hanish et al., 2004; McCulloch, Wiggins, Joshi, & Sachdev, 2000). It is possible that, because teachers do not expect externalizing behavior problems in females, the females' behavior problems are more salient and therefore are reported as more severe. Additionally, the sample included in this study was mostly male (a reflection of the ratio of males to females in the target population), so conclusions about females should be interpreted cautiously. Prior research has also shown that teacher reports may be subject to some bias when dealing with students with special needs (e.g. De Los Reyes & Kazdin, 2005), as demonstrated here by the finding that children in self-contained classrooms received lower *rule-breaking* ratings than those in regular classrooms.

However, the observation reports should not be considered a perfect index of behavior either. It is important to remember that, no matter how reliable, the observation data reflects only a small slice of behavior which may or may not be valid for predicting cognitive processes (such as perceived self-efficacy) that operate on a more general scale. Additionally, as was noted in the methods section above, this observation data set was not designed to capture student behavior, but to capture the interaction between student and teacher. As such, the current results, which use only a piece of this data, should be interpreted with caution.

The perceived self-efficacy measure used in the current study adds to existing literature in that it makes explicit the emotional aspects of the hypothetical situations on which it is based. However, the current study has highlighted several limitations to the perceived self-efficacy instrument. First, more scenarios and a widening of response options (e.g., from a 3 point scale to a 5 point scale) are needed to increase the variability and create a more continuous, normalized distribution. Second, the choice of scenarios should be reevaluated. A revised perceived self-efficacy measure should better align the context of the vignettes with the context of the behaviors reported through the CBCL-TRF and observation. Additionally, the perceived self-efficacy measure should be revised to eliminate the possible confound created by including both negative and positive emotions (see Harter & Buddin, 1987 for an example of how differently valenced emotions may be differently understood by children). Finally, because it is based on student interview, the perceived self-efficacy measure is susceptible to children misunderstanding the questions. The addition of a check for understanding would make future results clearer.

An additional measurement issue was the misalignment between the perceived self-efficacy, teacher-report, and observation measures, in both context and the types of behaviors on which they were focused. As was discussed above, it is this misalignment that may have contributed most to the muddled results. This is especially true for the relationship between perceived self-efficacy and the rule-breaking teacher-report scale, but not the aggression teacher-report scale or the observer reports of disruptive behavior.

Finally, because the current study included only children with high levels of externalizing behavior problems, it is possible that the restricted range of CBCL-TRF and observation scores affected the results. A restricted range is likely to result in artificially lowered beta values (Cohen, Cohen, West, & Aiken, 2003), so a follow-up study with a wider range of participants would be needed to increase confidence in the results.

Generalizability of Results

Finally, it should be noted that the current study includes only a limited age range from one geographical area, and that it includes only children with or at risk for EBD. Thus, much more work is needed before findings can be generalized to a typical school population. Additionally, the current results are correlational and cross-sectional, and so causal inferences cannot be drawn (see Shadish, Cook, & Campbell, 2002 for a discussion of the experimental designs under which results can be generalized and causal inferences drawn).

Future Directions

The immediate goal for this line of research should be to attempt to replicate these findings with more statistically sound, better aligned measures, and with a more diverse sample. Specifically, revising the perceived self-efficacy measure as described above is an important next step. In addition, the speculative theories presented above for explaining the results should be empirically evaluated. Specifically, the possible bias in the CBCL-TRF, the misalignment of measures, and the hypothesis that this high-risk sample is overestimating their competence should be tested as possible causes of the observed disconnect between perceived self-efficacy to control behavior and actual teacher/observer-reported behavior.

Also, the effect of speech/language delay on perceived self-efficacy and the mechanisms by which the effects operate should be examined to determine if this effect was a result of children misunderstanding the measure (it has been shown that children with speech or language delay often also exhibit difficulties with comprehension; e.g. Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998) or of a true difference in how these children perceive their emotionally-driven behavior.

Additionally, future study might include examination of the mediating role of Social Information Processing (SIP) within the relationship between perceived self-efficacy and externalizing behavior problems. The discrepancy between the current findings and those of Perry et al. (1986) and Gotthiel and Dubow (2001) suggest that children may reason differently about their SIP skills when they are explicitly reminded that successful processing requires the regulation of emotionally-driven responses.

Further study examining children's social problem-solving when emotion is made salient is needed to better describe this phenomenon.

Finally, the current study also raises questions about whether perceived self-efficacy can be considered a separate construct from self-reported ability, and the implications for research if the two are indeed indistinguishable. In an unpublished data set, Eisenberg and colleagues found a positive relationship between self-reported behavioral regulation and maladaptive social behavior among fourth and fifth grade boys (personal communication, 2009). These results, combined with the current results, bring into question the validity of measuring behavioral regulation with self-report measures, especially among children who are younger and/or show high levels of behavior problems.

If continuing research supports a role for perceived self-efficacy to control behavior in understanding problem behaviors, the implications for intervention should not be ignored. Current cognitive training for social skills currently stresses skills needed to competently process social problems, but results are mixed at best (see Magee Quinn, Kavale, Mathur, Rutherford, & Forness, 1999 for a review). If children are, in fact, overestimating their ability to regulate emotional responses, the inclusion of training to help children realistically gauge their success at regulation could be a valuable part of this training. Only continued research will determine the viability of this line of inquiry.

Appendix A: Perceived Self-Efficacy Vignettes

You are playing a game with your friend. At the end of the game, you win and your friend loses. You feel very happy and want to cheer. But you know that if you cheer, your friend will not want to play anymore.

You want to go outside, but your Mom says you can't because it is raining. You feel very upset and want to scream. Your friend is coming over to play. You know that your friend will not want to play with you if you are screaming.

Appendix B: Student Behavior MOOSES Codes Used for This Analysis

Negative talk: Score this for statements or vocalizations made with the intent to provoke, annoy, pester, mock, whine, complain, tattle, or make fun of another. This category also includes threats of physical aggression against person or property, arguing or disagreeing with another person (as in protest) as well as any verbal refusal to comply with a command. Code negative talk separately if at least five seconds have passed between **the end** of one statement and the beginning of the next.

Examples:
"Your mother."
"I'm going to kill you."
Protests: "Oh, man!" or "Hey, that's not fair!"
"No, I won't do it."
Curse words

Aggression: This is scored for deliberate physical contact that is potentially harmful to self, others, or property, and also for posturing or a gesture that is intended to provoke another. Code aggression separately if at least five seconds have passed between **the end** of one incident and the beginning of the next.

Examples:
Hitting, pushing, biting, kicking, or grabbing
Pulling someone's clothes
Throwing something at someone else
Shooting a bird

Non-Examples:
Swearing at another person
Accidentally bumping into another person
Putting both arms around someone else and hugging them

REFERENCES

28th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, vol. 2. (2009). Washington, DC: U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Program.

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

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anjum, R. (2006). The impact of self-efficacy on mathematics achievement of primary school children. *Pakistan Journal of Psychological Research, 21*, 61-78.

Asher, S. R., & Paquette, J. A. (2003). Loneliness and peer relations in childhood. *Current Directions in Psychological Science, 12*, 75-78.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Bandura, A., Pastorelli, C., Barbaranelli, C., & Caprara, G. V. (1999). Self-efficacy pathways to childhood depression. *Journal of Personality and Social Psychology, 76*, 258-269.

Caldarella, P., Young, E. L., Richardson, M. J., Young, B. J., & Young, K. R. (2008). Validation of the Systematic Screening for Behavior Disorders in middle and junior high school. *Journal of Emotional and Behavioral Disorders, 16*, 105-117.

Calkins, S. D., Gill, K. L., Johnson, M. C., & Smith, C. L. (1999). Emotional reactivity and emotional regulation strategies as predictors of social behavior with peers during toddlerhood. *Social Development, 8*, 310-334.

Camodeca, M., & Goossens, F. A. (2005). Aggression, social cognitions, anger and sadness in bullies and victims. *Journal of Child Psychology and Psychiatry, 46*, 186-197.

Campbell, S. B., Shaw, D. S., & Gilliom, M. (2000). Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. *Development and Psychopathology, 12*, 467-488.

Carter, E. W., Sisco, L. G., Brown, L., Brickham, D., & Al-Khabbaz, Z. A. (2008). Peer interactions and academic engagement of youth with developmental disabilities in inclusive middle and high school classrooms. *American Journal on Mental Retardation, 113*, 479-494.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, MJ: Erlbaum.

Coie, J. D., Dodge, K. A., Terry, R., & Wright, V. (1991). The role of aggression in peer relations: An analysis of aggression episodes in boys' play groups. *Child Development, 62*, 812-826.

Coie, J. D., Lochman, J. E., Terry, R., & Hyman, C. (1992). Predicting early adolescent disorder from childhood aggression and peer rejection. *Journal of Consulting and Clinical Psychology, 60*, 783-792.

Cole, P. M., Zahn Waxler, C., Fox, N. A., Usher, B. A., & Welsh, J. D. (1996). Individual differences in emotion regulation and behavior problems in preschool children. *Journal of Abnormal Psychology, 105*, 518-529.

Coleman, P. K. (2003). Perceptions of parent-child attachment, social self-efficacy, and peer relationships in middle childhood. *Infant and Child Development, 12*, 351-368.

Connolly, J. (1989). Social self-efficacy in adolescence: Relations with self-concept, social adjustment, and mental health. *Canadian Journal of Behavioral Science, 21*, 258-269.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74-101.

Crick, N. R., Grotpeter, J. K., & Bigbee, M. A. (2002). Relationally and physically aggressive children's intent attributions and feelings of distress for relational and instrumental peer provocations. *Child Development, 73*, 1134-1142.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483-509.

DiClemmente, C. C. (1981). Self-efficacy and smoking cessation maintenance: A preliminary report. *Cognitive Therapy and Research, 5*, 175-187.

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development, 51*, 162-170.

Dodge, K. A. (1983). Behavioral antecedents of peer social status. *Child Development, 54*, 1386-1399.

Dodge, K. A., & Price, J. M. (1994). On the relation between social information processing and socially competent behavior in early school-aged children. *Child Development, 65*, 1385-1897.

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

Eisenberg, N., Fabes, R. A., Bernzweig, J., Karbon, M., Poulin, R., & Hanish, L. (1993). The relations of emotionality and regulation to preschoolers' social skills and sociometric status. *Child Development, 64*, 1418-1438.

Eisenberg, N., Fabes, R. A., Guthrie, I. K., & Reiser, M. (2000). Dispositional emotionality and regulation: Their role in predicting quality of social functioning. *Journal of Personality and Social Psychology, 78*, 136-157.

Eisenberg, N., Fabes, R. A., Murphy, B., Maszk, P., Smith, M., & Karbon, M.. (1995). The role of emotionality and regulation in children's social functioning: A longitudinal study. *Child Development, 66*, 1360-1384.

Eisenberg, N., Guthrie, I. K., Fabes, R. A., Shepard, S., Losoya, S., Murphy, B. C., et al. (2000). Prediction of elementary school children's externalizing problem behaviors from attention and behavioral regulation and negative emotionality. *Child Development, 71*, 1367-1382.

Fox, J., & Davis, C. (2005). Functional behavior assessment in schools: Current research findings and future directions. *Journal of Behavioral Education, 14*, 1-4.

Gottheil, N. F., & Dubow, E. F. (2001). Tripartite beliefs models of bully and victim behavior. *Journal of Emotional Abuse, 2*, 25-47.

Hanish, L. D., Eisenberg, N., Fabes, R. A., Spinrad, T. L., Ryan, P., & Schmidt, S. (2004). The expression and regulation of negative emotions: Risk factors for young children's peer victimization. *Development and Psychopathology, 16*, 335-353.

Harris, P. L. (1989). *Children and emotion*. Oxford: Blackwell.

Harris, P. L., & Gross, D. (1988). Children's understanding of real and apparent emotion. In D. R. Olson, J. W. Astington, & P. L. Harris, Paul L (Eds.). *Developing theories of mind.* (295-314*).* New York: Cambridge University Press.

Harter, S., & Buddin, B. J. (1987). Children's understanding of the simultaneity of two emotions: A five-stage developmental acquisition sequence. *Developmental Psychology, 23*, 388-399.

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

Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology, 20*, 1120-1134.

Hymel, S., Bowker, A., & Woody, E. (1993). Aggressive versus withdrawn unpopular children: Variations in peer and self-perceptions in multiple domains. *Child Development, 64*, 879-896.

Hymel, S., Rubin, K. H., Rowden, L., & LeMare, L. (1990). Children's peer relationships: Longitudinal prediction of internalizing and externalizing problems from middle to late childhood. *Child Development, 61*, 2004-2021.

Individuals with Disabilities Education Act (2004). Retrieved May 11, 2009: http://idea.ed.gov/download/statute.html.

Kauffman, J. M., & Landrum, T. J. (2009). *Characteristics of emotional and behavioral disorders of children and youth* (9th ed.). Upper Saddle River, NJ: Merrill.

Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence* (2nd ed.). Newbury Park, CA: Sage.

Kupersmidt, J. B., & Coie, J. D. (1990). Preadolescent peer status, aggression, and school adjustment as predictors of externalizing problems in adolescence. *Child Development, 61*, 1350-1362.

Lee, C. (1984). Accuracy of efficacy and outcome expectations in predicting performance in a simulated assertiveness task. *Cognitive Therapy and Research, 8*, 37-48.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development, 71*, 107-118.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology, 31*, 356-362.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology, 33*, 265-269.

Locke, E. A., Frederice, E., Lee, C., & Bobko, P. (1984). Effect of self-efficacy, goals, and task strategies on task performance. *Journal of Applied Psychology, 69*, 241-251.

Magee Quinn, M., Kavale, K. A., Mathur, S. R., Rutherford, R. B., Jr., & Forness, S. R. (1999). A meta-analysis of social skill interventions for students with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders, 7*, 54-64.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

McCulloch, A., Wiggins, R. D., Joshi, H. E., & Sachdev, D. (2000). Internalising and externalising children's behaviour problems in Britain and the US: Relationships to family resources. *Children and Society, 14*, 368-383.

McDowell, D. J., O'Neil, R., & Parke, R. D. (2000). Display rule application in a disappointing situation and children's emotional reactivity: Relations with social competence. *Merrill Palmer Quarterly, 46*, 306-324.

Meier, S., McCarthy, P. R., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research, 8*, 107-120.

Merrell, K. W. (2001). Assessment of children's social skills: Recent developments, best practices, and new directions. *Exceptionality, 9*, 3-18.

Norwich, B. (1987). Self-efficacy and mathematics achievement: A study of their relation. *Journal of Educational Psychology, 79*, 384-387.

Ollendick, T. H., Weist, M. D., Borden, M. C., & Greene, R. W. (1992). Sociometric status and academic, behavioral, and psychological adjustment: A five-year longitudinal study. *Journal of Consulting and Clinical Psychology, 60*, 80-87.

Orobio de Castro, B., Bosch, J. D., Veerman, J. W., & Koops, W. (2003). The effects of emotion regulation, attribution, and delay prompts on aggressive boys' social problem solving. *Cognitive Therapy and Research, 27*, 153-166.

Orobio de Castro, B., Slot, N. W., Bosch, J. D., Koops, W., & Veerman, J. W. (2003). Negative feelings exacerbate hostile attributions of intent in highly aggressive boys. *Journal of Clinical Child and Adolescent Psychology, 32*, 56-65.

Paris, S. G., & Oka, E. R. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review, 6*, 25-56.

Perry, D. G., Perry, L. C., & Rasmussen, P. (1986). Cognitive social learning mediators of aggression. *Child Development, 57*, 700-711.

Randhawa, B. S., Beamer, J. E., & Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. *Journal of Educational Psychology, 85*, 41-48.

Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behavior. *Child Development, 55*, 1504-1513.

Samson, J. E., & Giles, J. W. (manuscript in preparation). Children's perceptions of the malleability of mental states versus behavior.

Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review, 1*, 173-208.

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

Shores, R. E., & Wehby, J. H. (1999). Analyzing the classroom social behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 7*, 194-199.

Skiba, R. J., Poloni-Staudinger, L., Gallini, S., Simmons, A. B., & Feggins-Azziz, R. (2006). Disparate access: The disproportionality of African American students with disabilities across educational environments. *Exceptional Children, 72*, 411-424.

Stata annotated output: Zero-inflated negative binomial regression. *UCLA Academic Technology Services* Retrieved May 5, 2009, from http://www.ats.ucla.edu/stat/stata/output/Stata_zinb.htm.

Stata data analysis examples: Zero-inflated negative binomial regression. *UCLA Academic Technology Services* Retrieved May 5, 2009, from https://www.ats.ucla.edu/stat/stata/dae/zinb.htm.

Stipek, D. J. (1993). *Motivation to learn: From theory to practice* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41*, 407-418.

Tapp, J., Wehby, J., & Ellis, D. A. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments, & Computers, 27*, 25-31.

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the Society for Research in Child Development, 59*, 25-52, 250-283.

Urbain, E. S., & Kendall, P. C. (1980). Review of social-cognitive problem-solving interventions with children. *Psychological Bulletin, 88*, 109-143.

VanOostrum, N., & Horvath, P. (1997). The effects of hostile attribution on adolescents' aggressive responses to social situations. *Canadian Journal of School Psychology, 13*, 48-59.

Walden, T., & Smith, M. C. (1997). Emotion regulation. *Motivation and Emotion, 21*, 7-25.

Walker, H. M., & Severson, H. (1990). *Systematic screening for behavioral disorders*. Longmont, CO: Sopris West.

Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology, 71*, 136-149.