Interpersonal Skills Group – Corrections Modified for Detained Juvenile Offenders with Externalizing Disorders: A Controlled Pilot Clinical Trial

A dissertation presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Nora Bunford

August 2016

This dissertation titled

Interpersonal Skills Group – Corrections Modified for Detained Juvenile Offenders with Externalizing Disorders: A Controlled Pilot Clinical Trial

by

NORA BUNFORD

has been approved for

the Department of Psychology

and the College of Arts and Sciences by

Steven W. Evans

Professor of Psychology

Robert Frank

Dean, College of Arts and Sciences

ABSTRACT

BUNFORD, NORA, Ph.D., August 2016, Clinical Psychology

Interpersonal Skills Group – Corrections Modified for Detained Juvenile Offenders with Externalizing Disorders: A Controlled Pilot Clinical Trial

Director of Dissertation: Steven W. Evans

The symptoms and correlates of externalizing disorders place youth with those disorders at-risk for criminal offending. Indeed, externalizing disorders are among the most common psychiatric disorders among detained juvenile offenders. Thus, effective treatments, that are appropriate for both the population and for delivery in juvenile detention, are needed. Yet, the state of the pertinent science is in its early stages and with limitations. Some limitations are methodological and some are clinical. Of import, the tested treatments do not systematically and simultaneously target emotion dysregulation and social impairment, despite basic findings indicating that both are associated with externalizing disorders and confer risk for criminal offending. To begin addressing these limitations, the purpose of the current study was to examine participant satisfaction with the Interpersonal Skills Group – Corrections Modified (ISG-CM) as well as the preliminary effectiveness of ISG-CM, in a controlled trial conducted at a juvenile detention facility and using multi-method and multi-informant measurement. Twelve detained juvenile offender youth (100% male; M_{age} = 16.30, SD_{age} = 1.16) participated. Results indicated that youth found the treatment highly satisfactory and that they found each of three treatment elements likeable and beneficial/helpful. Contrary to expectations, results were mixed with regard to changes in emotion regulation and social functioning.

In line with expectations, ISG-CM, relative to no treatment, was associated with either an attenuation of an increase in or with a decrease in self- and staff-rated verbal aggression, staff-rated aggression against property, and self-rated anger, across analyses. When in ISG-CM, relative to when not in treatment, youth also exhibited an increase in daily behavior points and those in in ISG-CM had fewer unsuccessful days in this domain than those not in treatment. Taken together, these results indicated that ISG-CM may be a promising approach to the psychosocial treatment of detained juvenile offenders with externalizing disorders. Clinical and methodological implications as well as testable hypotheses for future research on treatments for detained juvenile offenders are discussed.

DEDICATION

For my mentor and role-model in academia and humanity,

Steven W. Evans

and

For my eternal source of hope, strength, and wisdom,

my Mother

If you can dream and not make dreams your
master;
If you can think and not make thoughts your aim;
If you can meet with Triumph and Disaster
And treat those two impostors just the same;
[…]
If you can make one heap of all your winnings
And risk it on one turn of pitch-and-toss,
And lose, and start again at your beginnings
And never breathe a word about your loss;
If you can force your heart and nerve and sinew
To serve your turn long after they are gone,
And so hold on when there is nothing in you
Except the Will which says to them: 'Hold on!'
If you can talk with crowds and keep your virtue,
Or walk with Kings—nor lose the common
touch,
If neither foes nor loving friends can hurt you,
If all men count with you, but none too much;
[…]
Yours is the Earth and everything that's in it,
And—which is more—you'll be a Man, my son!

ha álmodol – s nem zsarnokod az álmod,
gondolkodol – becsülöd a valót,
ha a Sikert, Kudarcot bátran állod,
ugy nézed őket, mint két rongy csalót,
[…]
ha mind, amit csak nyertél, egy halomban,
van merszed egy kártyára tenni föl,
s ha vesztesz és elkezded újra, nyomban,
nem is beszélsz a veszteség felől,
ha paskolod izmod, inad a célhoz
és szíved is, mely nem a hajdani,
mégis kitartasz, bár mi sem acéloz,
csak Akaratod int: „Kitartani",
ha szólsz a néphez s tisztesség a vérted,
királyokkal jársz, s józan az eszed,
ha ellenség, de jóbarát se sérthet,
s mindenki számol egy kicsit veled,
[…]
tiéd a Föld és minden, ami rajta,
és – ami több – ember leszel, fiam.

- Rudyard Kipling

ACKNOWLEDGEMENTS

I am utterly grateful to Dr. Steven W. Evans for his unwavering encouragement, guidance, and support throughout my graduate studies.

I thank Dr. Peggy M. Zoccola for her enthusiasm towards pedagogy and time spent on introducing me to psychophysiology and on training me on the use of corresponding measurements.

I am extending a note of appreciation to members of my dissertation committee, including Drs. Julie Sarno Owens, Julie A. Suhr, and Gillian Ice.

I would like to acknowledge and thank Troy Thieman, Jennifer Conkey, Karly Dahn, and Shelby Kerwin for their dedication to this research and hard work.

I thank the youth who participated in this research for their participation and for reminding me that, often, "*Out of suffering emerge the strongest souls; The most massive characters are often seared with scars*".

I thank the Perry Multi County Juvenile Facility staff for their collaboration, especially Steve Koenig and Christopher Johnson.

I thank my family for their love and support and for making it possible for me to pursue my education and dreams.

This study was funded by the Elizabeth Munsterberg Koppitz Fellowship awarded by the American Psychological Foundation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Attention-Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD) (hereafter: externalizing disorders[1]), are prevalent, impairing, confer risk for criminal offending, and are costly to individuals and society. Emotion dysregulation (ED) and social impairment are two characteristics that are both associated with externalizing disorders (e.g., Beauchaine, Gatzke-Kopp, & Mead, 2007; Bunford, Evans, & Langberg, 2014; Greene et al., 2002; Nelson-Gray, et al., 2006; Walker et al., 1991) and confer risk for criminal offending (e.g., Davidson, Putnam, & Larson, 2000; Hirschi, 1969). Although many youth with externalizing disorders are arrested and detained, the science on treatments that are adapted for delivery in juvenile detention settings is in its early stages (see Desai et al., 2006 for review) and the pertinent research completed to date has limitations. The limitations are methodological as well as clinical insofar as the treatments do not systematically target both ED and social impairment. This latter limitation is problematic given evidence indicating that ED and social impairment are associated with externalizing disorders and confer risk for criminal offending. Given these considerations, there is a need to develop treatments for detained juvenile offenders that target both ED and social impairment and to evaluate those treatments in methodologically sound research. In line with the steps outlined in the Deployment-Focused Model of Intervention Development (DFM; Weisz, Jensen, & McLeod, 2005) the goal in the current study was *to examine participant satisfaction* with

[1] Although the externalizing spectrum disorders include ADHD, ODD, CD, antisocial personality disorder (ASPD), pyromania, kleptomania, intermittent explosive disorder (IED), and substance-related disorders, the focus in the current study was on those that occur in youth and those that occur most frequently, i.e., ADHD, ODD, and CD.

the Interpersonal Skills Group – Corrections Modified (ISG-CM) as well as *the preliminary effectiveness* of ISG-CM in improving ED, social impairment, and functioning in related domains*, in a controlled pilot clinical trial* conducted at a juvenile detention facility and using multi-method and multi-informant measurement.

Externalizing Disorders are Costly, Impairing, and Confer Risk for Criminal Offending

Prevalence estimates of ADHD in the community range from 1.7% to 17.8%, of ODD from 2.8% to 5.5%, and of CD from 2.0% to 3.3% (Merikangas, Nakamura, & Kessler, 2009). The correlates of these disorders include aggression and cruelty to animals and people; defiance; difficulties with concentration and hyperactivity; deceitfulness or theft; destruction of property; violation of legal and social rules; and social impairment (American Psychiatric Association [APA], 2013). These correlates confer risk for criminal offending (e.g., Pulkkinen & Pitkanen, 1993) and many youth with externalizing disorders are arrested and detained. Indeed, ADHD, ODD, and CD are among the most common psychiatric disorders among detained juvenile offenders (Teplin et al., 2002). Among detained males and females, respectively, 16.6% and 21.4% have ADHD, 14.5% and 17.5% ODD, and 37.8% and 40.6% CD (Teplin et al., 2002).

Externalizing disorders and their correlates are costly to youth and society both independent of detention and also due to correlates of detention (e.g., increased recidivism and psychiatric symptom severity: Kirchner, 2014; reduced employment or occupational success after release: Holman & Zeidenburg, 2013). Related, the average financial cost of juvenile detention in the U.S. is, per detained youth, $407 a day, or $148,767 per year (Kirchner, 2014).

Taken together, these data indicate that externalizing disorders are prevalent, impairing, confer risk for criminal offending, and are costly. Given these impairments, risk, and costs, identifying characteristics that contribute to their development and maintenance is key for prevention and treatment. Prior theory and findings indicate that ED and social impairment are two such characteristics.

ED, Social Impairment, and Externalizing Disorders

ED is associated with ADHD, ODD, and CD (e.g., Beauchaine et al., 2007; Bunford et al., 2014; Bunford, Evans, Becker & Langberg, 2015; Hinshaw, 2003; Nelson-Gray, et al., 2006) and with behaviors that confer risk for criminal offending (e.g., aggression and violence; Bushman et al., 2001; Calkins, Gill, Johnson, & Smith, 1999; Davidson et al., 2000; Furlong & Smith, 1994 as well as alcohol/other drug abuse: Bradley et al., 2011). Social impairment is also associated with externalizing disorders (e.g., ADHD: Bunford et al., 2014, ODD: Greene et al., 2002, CD: Walker et al., 1991) and confers risks for criminal offending in part through association with deviant peers (Brezina & Piquero, 2007; Chen & Adams, 2010; Hirschi, 1969; Owens & Slocum, 2015). Thus, there is evidence that ED and social impairment are associated with externalizing disorders and place youth with those disorders at risk for criminal offending. Given that externalizing disorders place youth at risk for arrests and detention and the high prevalence of externalizing disorders in detention, it is prudent that treatments for these disorders and correlates are available in juvenile detention facilities. Indeed, one purpose of incarceration is rehabilitation, that is, to begin addressing the difficulties that contribute to offenders' transgressions and thereby reduce recidivism

(e.g., Dembo et al., 1977; McCord, Widom, & Crowell, 2001; Timmons-Mitchell et al., 1997).

The Science on Treatments in Juvenile Detention is Limited

Despite the need for available treatments in detention facilities and rehabilitation being a purpose of detention, there is a paucity of treatments that have been adapted for delivery in juvenile detention facilities. In addition, there is even less empirical data *obtained in juvenile detention settings* on the effectiveness of such treatments (Desai et al., 2006). Exceptions to these limitations are half a dozen studies with encouraging findings. The techniques tested in these studies target ED and interpersonal or social skills. The authors of four non-randomized studies tested the effects of dialectical behavior therapy (DBT; Shelton et al., 2011, Trupin et al., 2002), cognitive restructuring of pro-aggression thoughts (Guerra & Slaby, 1990), and behavior modification and peer counseling (Ross & McKay, 1976). The authors of three randomized controlled trials tested social skills training (Marziller & Spence, 1981), stress inoculation training (Schlicter & Horan, 1981), and a treatment targeting interpersonal or social behaviors (Shivrattan, 1988).

Of the studies wherein beneficial effects were observed, limitations include non-random assignment (Ross & McKay, 1976; Shelton et al., 2011; Trupin et al., 2002), pre-treatment differences among groups that prevented between-group comparison (Trupin et al., 2002), and measurement either of only one (i.e., aggression; Guerra & Slaby, 1990) or only a few (Schlichter & Horan, 1981; Shivrattan, 1988) of the many outcomes relevant to the functioning of detained juvenile offenders.

The treatment components that appear to be promising across the tested treatments include those in which adolescents receive feedback from peers (Ross & McKay, 1976) and that youth are exposed (either via imagery or in vivo via simulated provocations) to actual social situations (Marziller & Spence, 1981; Shivrattan, 1988). In the majority of these studies, counts and percentages are reported as indices of change (e.g., number of antisocial acts or percent of youth who return to the facility after release). Effect size estimates are not reported in any of the studies and neither are *M* or *SD* values that would allow for the calculation of such estimates. One exception is the study by Schlichter and Horan (1981) who reported *M* and *SD* values corresponding to their outcome measurements. To estimate the magnitude of the observed effects based on the data these authors reported, repeated measures effect sizes were calculated, in line with Becker (1988) using the following formula: $((M_{\text{post, treatment group}} - M_{\text{pre, treatment group}})/SD_{\text{pre, treatment group}}) - (M_{\text{post, waitlist group}} - M_{\text{pre, waitlist group}})/SD_{\text{pre, waitlist group}}))$. In response to audiotaped anger provoking events, youths' response with regards to how angry they would become if they were actually involved in each situation indicated post-treatment improvements for both stress inoculation training and the treatment elements condition (including some but not all components of stress inoculation training) relative to no treatment control, with a large- and a medium-sized effect, respectively (d = 1.10 and 0.66). In an experimental paradigm, each youth interacted with a research staff member unaware of treatment condition and engaged in a role-play of provocative events. The ratings of independent judges unaware of treatment condition indicated post-treatment improvements in verbal aggression for both stress inoculation training and treatment

elements relative to no treatment control, with a large- and a medium-sized effect, respectively (d = 0.97 and 0.59). Finally, the ratings of facility staff unaware of treatment condition indicated post-treatment improvements for both stress inoculation training and treatment elements relative to no treatment control on both verbal and physical aggression. The corresponding repeated measures effect sizes were small (d = 0.29) and medium (d = 0.49) (for stress inoculation and treatment elements, respectively) for verbal aggression and were large (d = 1.26) and large (d = 1.53) for physical aggression (for stress inoculation and treatment elements, respectively).

In addition to feedback from peers and practice in realistic social situations, DBT (a treatment that in essence is a comprehensive treatment for ED; Linehan, Bohus, & Lynch, 2007) is also supported by some evidence, highlighting the additional importance of targeting ED. Taken together, the literature reviewed hitherto indicates the need to (1) develop a treatment for detained juvenile offenders with externalizing disorders that targets both social impairment and ED and (2) test this treatment in the context of a controlled trial, measuring multiple indices of outcomes.

Rationale for Interpersonal Skills Group – Corrections Modified

To this end, the Interpersonal Skills Group (Evans et al., 2016; Evans, Schultz, & DeMars, 2014; Kern, Evans, Lewis, Weist, & Wills, 2015) was chosen as a starting point for the present treatment development and evaluation process. Although ISG does not meet criteria for well-established treatment (Evans, Owens & Bunford, 2014), there is more evidence supporting its effects in addressing social impairment associated with ADHD and related externalizing problems among adolescents than any alternatives (e.g.,

Sadler, Evans, Schultz & Zoromski, 2011). Standard ISG is comprised of three phases, Psychoeducation (Phase I), Skill Building (Phase II), and Generalization (Phase III). ISG involves (1) adolescents establishing goals about how they would like to be perceived by others (Ideal Self Goals) and identifying behaviors that are congruent with those goals/perceptions and (2) training adolescents to consider nonverbal and verbal social feedback and modify their behaviors given their Ideal Self Goals and social feedback. ISG was chosen to begin the present treatment modification and evaluation process given (1) the relevance of its theoretical underpinnings to detained juvenile offenders and (2) its applicability to adolescents.

First, the groundwork for ISG is based on developmental theory and social cognitive data on deficits that contribute to social impairment. The social cognitive deficits associated with ADHD include deficits in the ability to establish and understand cause and effect relationships in social situations (e.g., Lorch, Milich, Astrin, & Berthiaume, 2006). The social cognitive deficits associated with aggression (which youth with ADHD, ODD, and CD exhibit more frequently than youth without these disorders) involve deficits in social problem solving (Lochman & Lampron, 1986; Slaby & Guerra, 1988). Among youth with ADHD (Sibley, Evans, & Serpell, 2010) and among youth with aggression (Schultz, Izard, & Ackerman, 2000), these respective social cognitive deficits are associated with social impairment. In ISG, both types of social cognitive deficits are directly targeted.

Second, ISG is a training intervention (TI) – an intervention wherein the main focus is on training without manipulating contingencies in the environments where

behavior change is intended to occur – and the treatment model of TIs has been conceptualized as the preferred model of treatment for adolescents (Evans et al., 2014b). Because TIs do not involve manipulating contingencies in the environments where behavior change is intended to occur[2], there is a corresponding assumption that TIs produce changes that persist across settings and over time (Evans et al., 2014b). There is evidence that ISG results in long-term improvements (Schultz, Evans, Langberg & Schoemann, in preparation) that foster autonomy and independence (key developmental tasks of adolescence). This makes ISG particularly appropriate for detained juvenile offenders as the competencies these youth learn in treatment in detention will need to be useable primarily in settings other than the treatment setting. Thus, given its focus on improving social cognitive deficits and that TIs are indicated for adolescents and likely result in changes that persist across settings and over time, it was hypothesized that ISG, modified for a detention setting, would be particularly appropriate for treating detained juvenile offenders.

However, ISG does not directly address ED, despite reasons to believe that ED is an important treatment target for detained juvenile offenders with externalizing disorders. Indeed, problems redolent of ED (i.e., parent-rated *Emotional Symptoms* on the Behavior Assessment System for Children) were associated with treatment response among 15 adolescents with ADHD (86.4% male, between 13 and 17 years of age; [Cohen's d = 0.56]) in a re-analysis (Bunford & Evans, 2013; unpublished manuscript) of data obtained in a previous study of ISG (Sadler et al., 2011). Thus, it was hypothesized that response

[2] Conversely to behavioral management treatments which are intended to lead to behavior change by manipulating contingencies in the target environment.

to ISG would be improved by directly targeting ED. To this end, the original ISG protocol of Evans and colleagues (Evans et al., 2016; Evans et al., 2014a; Kern et al., 2015) was modified to directly target ED, following the procedures described next.

Development of Interpersonal Skills Group – Corrections Modified

To modify the original ISG protocol, primarily, the treatment development and evaluation steps outlined in the DFM (Weisz et al., 2005) were followed. The first three steps include step one: developing and refining the treatment protocol; step two: piloting the manual with non-symptomatic individuals; and step three: conducting a case study to inform progressive adaptations to the manual. Because step one is to be grounded in clinical practice, theory, and empirical findings on the target condition and corresponding *mechanisms of change*, in keeping with the treatment development process indicated by the Institute for Education Sciences (IES; n.d.), the first step in the current development process involved identifying an impairing and malleable characteristic of externalizing disorders.

Step One: Identify Impairing and Malleable Characteristics of Externalizing Disorders

The results of a series of studies (Bunford et al., 2014; Bunford et al., 2015) indicated that three aspects of ED, 1) emotional inflexibility and a slow return to emotional baseline; 2) a low threshold for emotional excitability, impatience, and socially inappropriate behaviors; and, 3) difficulty in behavioral control when experiencing negative emotions, are uniquely (beyond ODD and CD) associated with social impairment among youth with ADHD. Other domains of ED, such as the extent to which youth experience difficulty exhibiting socially appropriate emotional responses, such as

empathy and friendliness, were uniquely (beyond ADHD) associated with ODD. Thus, these data indicated that ED is a socially impairing characteristic of ADHD and related externalizing problems.

Regarding malleability, although no pertinent research was conducted with youth with externalizing disorders, others have found that receipt of DBT was associated with improvements on rating scale measures of ED (e.g., Axelrod, Perepletchikova, Holtzman, & Sinha, 2011: among adults with borderline personality disorder; Goldstein, Axelson, Birmaher, & Brent, 2007: among adolescents with bipolar disorder; and Telch, Agras, & Linehan, 2000: among adults with binge-eating disorder). Thus, these data indicated that ED is a malleable characteristic.

Step Two: Manualizing and Refining the Treatment Protocol

Having identified ED as an impairing and malleable characteristic, modifications to the original ISG protocol to directly target aspects of ED associated with ADHD and related externalizing problems were based on clinical practice and the empirical literature. Specifically, psychoeducation was included about emotion regulation and its implications for developing and maintaining social relationships and mindfulness meditation. It was made a requirement that participants' Ideal Self Goals and their operational definitions are related to emotion regulation. Feedback sessions during practice activities were modified to include self-ratings of emotional arousal on an Emotions Thermometer (see Appendix) and a discussion of the link between this arousal and relative difficulty in achieving Ideal Self Goals. Practice activities were modified to systematically elicit increasingly higher levels of frustration and mindfulness mediation

practice was added (at the conclusion of every group session and encouraged to be practiced outside of group) (see Table 1 for modification details).

Step Three: Piloting the Manual with Non-Symptomatic Individuals

Per ethical reasons outlined in the DFM, two pilot studies were conducted with non-symptomatic young adults ($N = 9$) and adolescents ($N = 8$) to examine the perceived safety and subjective efficacy of the mindfulness meditation element and the Psychoeducation element. In both pilot studies, the treatment elements were delivered by a doctoral-level clinician who has had considerable prior experience in delivering manualized treatments, including the traditional ISG protocol. Supervision was provided by a licensed clinical child psychologist.

No participants reported negative consequences, supporting the perceived safety of the mindfulness meditation element and the Psychoeducation element. The young adults found mindfulness practice beneficial in terms of an increase in relaxation, supporting the subjective efficacy of this element. Because some young adults fell asleep during meditation and some reported that the practice was too long, the protocol was modified to include the original, 20-minute guided meditation and a new, seven-minute guided meditation. Roughly half of the adolescents reported that they found Psychoeducation beneficial in terms of learning new information about emotions and emotion regulation and goal setting (i.e., Ideal Self Goals), supporting the subjective efficacy of this element. Because some youth reported that they found the Psychoeducation element “boring” and “too long”, Phase I was modified by shortening time spent on verbal instruction and the inclusion of engaging activities to demonstrate

key concepts (e.g., to demonstrate the difficulty of redirecting attention away from an intense emotion, youth hold ice-cubes, are directed to focus on physical features of the ice, and then are directed to redirect their attention to mundane aspects of the treatment room and then reflect on the degree to which such redirection was easy/difficult).

Step Four: Case Study to Inform Adaptations to the Manual

Following modifications to the original manual, the revised protocol was piloted in a case study design with adolescents with externalizing disorders ($N = 3$) recruited from a local high school and local health and mental health practices. Youth were between ages 14-18 years old, all identified as Caucasian, and all had ADHD (one also had ODD). None had a condition that would interfere with their ability to participate or achieve gains in treatment (e.g., psychotic symptomology) and none were receiving outpatient psychotherapy. In this case study, all three treatment elements were piloted, including the shortened Psychoeducation phase, Phase II, and the two mindfulness meditation tracks. The treatment was delivered by the same doctoral-level clinician as in the earlier two pilot studies. In line with the delivery model of ISG employed in prior studies of the treatment (e.g., Evans et al., 2016), Phase II feedback sessions were conducted by this clinician as well as by an undergraduate student. Supervision was provided by a licensed clinical child psychologist. The undergraduate student received 9 hr of training prior to beginning the program.

Qualitative reports indicated that youth found the Psychoeducation element acceptable, highly enjoyed Phase II activities and one youth indicated he liked both mindfulness meditation practices and regularly practiced both, another youth indicated he

liked the shorter mindfulness meditation practice and regularly practiced it, and the third youth indicated he did not like either mindfulness meditation practice and did not regularly practice. In light of these reports, no further changes were made to the protocol. However, the two clinicians initially had difficulty with completing the feedbacks in a timely manner (i.e., within three to five minutes). Youth are pulled from group activities twice during each Phase II session for feedback. In the context of feedbacks, youth and clinicians rate the youth's behavior, discuss their ratings, and brainstorm ideas for achieving improvement (see *Weekly Ideal Self Goals ratings* below). It is important for clinicians to efficiently complete the feedbacks to ensure that youth have sufficient time to participate in group activities and achieve improvement. Clinical supervision thus included a direct focus on conducting feedbacks and feedbacks were randomly timed throughout the remainder of the pilot. By the final four Phase II sessions, clinicians were able to complete feedback sessions within the indicated amount of time.

Step Five: Group-Design Studies

The fifth step in the DFM involves a series of group-design studies testing the newly adapted treatment to explore, among others, the extent to which the treatment works with symptomatic youth, in clinical care settings, and when compared with usual care in the setting of interest.

Evaluation of Interpersonal Skills Group – Corrections Modified: Current Study

The current study initiates the fifth step of the DFM process and is the most recent step in the present line of treatment modification and evaluation research. The goals of the current study were to examine participant satisfaction with the Interpersonal Skills

Group – Corrections Modified (ISG-CM) and the preliminary effectiveness of ISG-CM, in a controlled pilot clinical trial conducted at a juvenile detention facility and using multi-method and multi-informant measurement. To this end, the aim was to address the following research questions: (1) Did youth find ISG-CM satisfactory? Which treatment elements did they like and which ones did they find beneficial/helpful? (2) Did ISG-CM result in decreased ED and social impairment on proximal measures? (3) Did ISG-CM result in decreased ED and social impairment on distal measures? (4) Did ISG-CM result in decreased aggression, ADHD and ODD symptoms, academic impairment, and externalizing/internalizing problems?

Regarding multi-method measurement, with ED being one of the primary outcomes of interest, in addition to rating scale measures of ED, a measure of autonomic nervous system activity as a physiological marker of ED, heart-rate variability (HRV) (Thayer, Åhs, Fredrikson, Sollers, & Wager, 2012) was included. Given that clinical measures are often inconsistent and weak predictors of treatment response (see Ball, Stein, & Paulus, 2014 for review), it is important to identify alternative measures of constructs of interest that may serve as indices of change in response to treatment or as predictors of treatment response, or both. HRV was chosen as it has been shown to be feasible to assess with youth with ADHD (Bunford, Evans, Zoccola, & Rybak, 2013) and because it is highly cost-effective. A task (an unsolvable find-a-word puzzle) previously used with children with ADHD to elicit frustration (Milich, 1994) was modified[3] so as to

[3] Modifications to the task involved alterations to ensure age-appropriateness; I changed the task to be more complex and frustrating. The modified task is comprised of a letter matrix and a 9-letter target word presented below the matrix (on a sheet of paper). One- through eight-letter snippets (more than one of each) of the target word are included in the letter matrix but the full 9-letter target word is not. To increase

utilize an experimental paradigm during which to measure HRV. Prior findings confirmed that the modified task elicited frustration and produced effects measurable by HRV among adolescents with ADHD (Bunford et al., 2013), providing support for the inclusion of HRV as a physiological index of ED in the current study.

---

motivation, youth are told before the task that most individuals their age are able to complete it within the allotted time. For a copy of or more information on the puzzle, contact the authors.

CHAPTER 2: METHOD

The current study was conducted at the Perry Multi-County Juvenile Facility (PMCJF). The minimum length of stay at the PMCJF is 18 weeks but most youth stay for multiple months. Among other factors, the length of stay is determined based on youths' progress with PMCJF programs. Given the length of ISG-CM (i.e., seven weeks), 24 youth who had at least 18 weeks left in their sentence were considered for this study. These 24 youth were first randomized to participate (n = 16) or not participate (n = 8) in the present research and then, those randomized to participate, were randomized to either ISG-CM (treatment group; n = 8) or to waitlist (waitlist control group; n = 8). Both randomization procedures were carried out through stratified random sampling[4]. This study was conducted in compliance with Part 46 of Title 45 of the Code of Federal Regulations (CFR), Subpart C, *Additional Protections Pertaining to Biomedical and Behavioral Research Involving Prisoners as Subjects* and was approved by the local Institutional Review Board.

Participants

Participants were 12 detained juvenile offenders (age range: 14-18 years [M = 16.30, SD = 1.16]). All were male and 60% identified as Caucasian, 20% as African American, 10% as American Indian or Alaskan, and 10% as Biracial/Multiracial. One youth reported symptoms and duration consistent with social anxiety disorder, three with

[4] This randomization method was chosen because the PMCJF detainees are housed in three units. Housing assignment is dependent on behavior problem severity and legal history (i.e., the three units are not equivalent with regard to behavior problem and symptom severity). Stratified random sampling allowed for relative homogeneity across youth participating in the study and youth not participating as well as across the intervention and waitlist groups with regard to behavior problem severity. The three units were treated as strata from which participants were selected via simple random sampling.

generalized anxiety disorder, three with post-traumatic stress disorder, four with major depressive disorder, and six with substance abuse (see Table 2). Seven youth self-reported symptoms, duration, and impairment consistent with a diagnosis of ADHD and the parents of six of these reported symptoms on the Disruptive Behavior Disorders Rating Scale (DBD; Pelham, Gnagy, Greenslade, & Milich, 1992) (see below) that reached diagnostic threshold for ADHD (data from one youth's parent was missing) (Table 3). Regarding ODD, seven youth self-reported symptoms, duration, and impairment consistent with a diagnosis and the parents of five of these reported symptoms on the DBD that reached diagnostic threshold (data from one youth's parent was missing) (Table 4). Twelve youth self-reported symptoms, duration, and impairment consistent with a diagnosis of CD and the parents of six of these reported symptoms on the DBD that reached diagnostic threshold (data from one youth's parent was missing) (Table 5). Youths' average two-subtest IQ score on the Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II; Wechsler, 1999) (see below) fell in the average range (M = 91.42, SD = 11.95, ranging from 67 to 108).

Youth randomly assigned to the present research underwent an eligibility evaluation. To adhere to ethical guidelines of the CFR, only one exclusionary criterion was applied. The exclusionary criterion was exhibiting symptoms of psychosis (delusions or hallucinations), determined via a semi-structured clinical interview (see below). This criterion was applied so that only youth who were likely to benefit from and fully participate in treatment were included. These procedures resulted in 16 eligible participants (i.e., no youth were excluded). Of these 16, a total of six dropped out of the

study due to unexpected discharge from the facility. Of the six who dropped out, four dropped out at the time ISG-CM was delivered to the first group. Three dropped out from the treatment group and one dropped out of the control group. Of the three who dropped out from the treatment group, one dropped out after the first session and was replaced by another youth (he was offered a makeup session; see below). The one who dropped out of the control group dropped out after mid-treatment and was replaced by another youth. Of the six who dropped out, two dropped out at the time ISG-CM was delivered to the second group. Both dropped out from the treatment group and none were replaced. The two replacement youth were not randomly selected; they were entered into the present research based on the PMCJF director's request. This process resulted in 12 youth who completed all study procedures and thus comprised the study sample. Effect size estimates were calculated to compare the treatment group pre-treatment means to the control group pre-treatment, in line with Hedges (1981; 1982) using the following formula: ($M_{\text{pre, treatment group}}$ - $M_{\text{pre, waitlist group}}$)/(SQRT(($SD_{\text{pre, treatment group}}$*$SD_{\text{pre, treatment group}}$+$SD_{\text{pre, waitlist group}}$*$SD_{\text{pre, waitlist group}}$)/2)). The treatment and waitlist groups did not differ with regard to age (Cohen's $D = 0.16$, with the first cohort being older to a negligible degree) or two-subtest IQ score (Cohen's $D = 0.09$, with the first cohort having higher IQ to a negligible degree).

Procedures

The design of the current study was a waitlist controlled design that began with eight participants in ISG-CM (i.e., treatment condition) and eight participants in control (but ultimately included six in each group). All 12 completed an initial evaluation (T1,

see below). Treatment began immediately after T1, lasted seven weeks, and was followed by post-treatment assessments (T2, see below). After the assessments, those initially in the control condition received the treatment for seven weeks, followed by a final evaluation (T3, see below). The Consort Diagram in Figure 1 depicts these procedures.

As noted, study assessments were conducted at three time-points and data on certain measures were collected throughout the duration of the study. The first assessment was the eligibility evaluation, which also served as the pre-treatment evaluation (T1). After the first group completed ISG-CM, all youth (those in ISG-CM and in waitlist) underwent a post-treatment evaluation (T2). A second post-treatment evaluation took place after the treatment for the second group was completed (T3). Data for these assessments was obtained from juvenile facility staff, juvenile facility teachers, youths' parents, and the youth.

Table 6 presents information on which measures were administered to which group and at which time-points. *T1 measurements* included a diagnostic interview and cognitive testing with all youth. HRV measurements were obtained from youth assigned to the first treatment group. All youth completed self-ratings of ED, social functioning, aggression, and externalizing/internalizing problems. Data was also collected from the parents of all youth. Teachers and staff completed ratings of ED, social functioning, and of aggression, ADHD and ODD symptoms, and academic functioning. *During the time that the first group was receiving the treatment*, staff completed ratings of ED and social functioning for youth in the treatment group only. As part of facility procedures, staff

also maintained a point system to manage youth behavior. Daily behavior points were available for youth in the treatment group and youth in the control group.

T2 measurements included HRV measurements obtained from all youth, with such measurements providing post-treatment data for the first treatment group and pre-treatment data for the second treatment group. All youth completed self-ratings of ED, social functioning, aggression, and externalizing/internalizing problems. Teachers and staff completed ratings of ED, social functioning, and of aggression, ADHD and ODD symptoms, and academic functioning. Youth in the first treatment group also completed ratings of treatment satisfaction and preferred treatment elements. *During the time of the second group receiving the treatment*, staff completed ratings of ED and social functioning for youth in the treatment group only. As part of facility procedures, staff also maintained a point system to manage youth behavior. Daily behavior points were available for youth in the treatment group (and for some youth from the first treatment group, however some youth from the first treatment group had been successfully discharged during the time of the second treatment group).

All *T3 measurements* were obtained from youth assigned to the second treatment group only (and were obtained for some youth from the first treatment group, however some youth from the first treatment group had been successfully discharged by the time of T3). T3 data obtained from youth from the first treatment group are not analyzed in the current study. Thus, for the sake of clarity and manageability, T3 will be discussed hereafter as involving assessments of the second treatment group only. T3 measurements included HRV measurements, self-ratings of ED, social functioning, aggression, and

externalizing/internalizing problems. Teachers and staff completed ratings of ED, social functioning, and of aggression, ADHD and ODD symptoms, and academic functioning. Youth also completed ratings of treatment satisfaction and preferred treatment elements.

Regarding treatment procedures, one group received ISG-CM first and those in the control group received ISG-CM second. Make-up sessions were offered to youth who missed the Phase I session. Because Phase II sessions necessitate a social context, no make-up sessions were offered to youth who missed ≤one session during Phase II. All evaluations and treatment were conducted at the PMCJF. In line with the delivery model of ISG employed in prior studies of the treatment (e.g., Evans et al., 2016), ISG-CM was staffed by a site supervisor (a doctoral-level clinician) who has had considerable prior experience in delivering manualized treatments, including the traditional ISG protocol. The site supervisor led group activities and supervised an undergraduate student. The undergraduate student had prior experience with delivering Phase II feedback sessions and received nine hours of training prior to beginning the program. The undergraduate student received 30 to 60 min of individual supervision weekly. The doctoral-level clinician received 30 min of individual supervision, weekly from a licensed clinical child psychologist.

The PMCJF, youth, and youths' parents received compensation for their participation. The PMCJF was compensated with gym equipment to be used by all detained youth ($500-value). All detained youth, regardless of their participation in this study, received consumables once every week (e.g., pizza and soft drinks) ($150-value per week). Youth who participated in this study received mp3 players (with two

mindfulness tracks recorded onto those, to encourage practice) ($45-value). Youth were able to sign in/out their mp3 players throughout the duration of the treatment. Although upon treatment completion youth turned the mp3 players in to PMCJF inventory, they could claim these upon their release. Parents received a $30 payment if they returned the rating scale measures.

Measures

Measures included a diagnostic interview, a test of cognitive ability, descriptive measures collected from parents at T1 only and proximal and distal outcome measures. All participants also completed ratings of satisfaction with the treatment and rated the treatment elements with regard to the degree to which those were likeable and beneficial/helpful. See description below and Table 6 for schedules for data collection.

Diagnostic Interview

Diagnoses were determined using the Children's Interview for Psychiatric Syndromes (ChIPS*;* Weller, Weller, Fristad, Rooney, & Schechter, 2000). The ChIPS is a semi-structured diagnostic interview of 20 DSM-IV-TR (APA, 2000) mental health disorders in youth 6-18 years. The ChIPS has adequate psychometric properties (Fristad et al., 1998; Weller et al., 2000). The ChIPS was administered to all youth at T1 by master's- and doctoral-level clinicians.

Cognitive Ability

Youths' cognitive ability was estimated using the Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II; Wechsler, 1999). The two-subtest form of the WASI-II (FSIQ-2) has adequate psychometric properties (Homack & Reynolds, 2007;

Wechsler, 1999). The two-subtest form was administered to all youth at T1 by master's- and doctoral-level clinicians.

ADHD, ODD, and CD Symptoms

The 45-item Disruptive Behavior Disorders Rating Scale – Parent Report (DBD; Pelham et al., 1992, see Appendix) is a measure of parent-rated ADHD, ODD and CD symptoms. Although there is limited psychometric data on the DBD–Parent, data on factorial validity (Pillow, Pelham, Hoza, Molina, & Stultz, 1998; Van Eck, Finney, & Evans 2010; Zuddas et al., 2006) and internal consistency (Van Eck et al., 2010) indicate that the DBD–Parent has adequate psychometric properties. Higher values indicate more frequent and severe symptoms. The DBD–Parent was mailed to all youths' parents at T1. Data from the DBD–Parent was used for descriptive purposes.

Proximal Outcome Measures

ISG-CM was designed to primarily improve ED and social functioning. To determine treatment effects on these outcomes, data on proximal measures of ED and social functioning was collected from both participating youth and from clinicians working with the youth in ISG-CM. The proximal measure of ED were HRV data collected prior to, during, and following youth engaging in the experimental frustration task. The proximal measure of social functioning was weekly self- and clinician-ratings of Ideal Self Goal mastery.

HRV Data Acquisition and Processing

HRV data was collected during a five-minute resting baseline (prior to youth working on the frustration task), during a five-minute frustration task, and during a five-

minute recovery (following youth working on the frustration task). HRV was recorded using a Polar Pro Trainer RS800 system (Polar Electro, Kempele, Finland), comprised of a wristwatch and a bio-harness. Youth wore the bio-harness on their upper torsos with the center of the harness located over the xiphisternum[5]. Data was recorded with Polar Pro-Trainer 5 software (PPT 5) (Polar Electro, Kempele, Finland). Data on variability in the time interval between heartbeats was extracted and corrected for artifacts using Kubios HRV Analysis Software 2.0 (Tarvainen, Niskanen, Lipponen, Ranta-Aho, & Karjalainen, 2014). Kubios HRV 2.0 was used to calculate the root mean square of successive differences (RMSSD) between adjacent time intervals (measured in msec) between heartbeats, to index HRV[6]. Higher values indicate less ED. The experimental frustration task described above (i.e., an unsolvable find-a-word puzzle) was used in the present study to elicit frustration. HRV data was collected at T1 and T2 for youth in the first treatment group and at T2 and T3 for youth in the second treatment group.

Weekly Ideal Self Goal Ratings

Self- and clinician ratings of Ideal Self Goal mastery occurred twice during each Phase II session. Youth rate their own behavior and clinicians observe youth and rate youths' behavior during group activities. Ratings are based both on youths' behavior and others' reactions to such behavior. Higher values indicate better social functioning (on a scale of "- 3" to "+ 3", "- 3" indicates behaviors that likely result in an impression by others that is incongruent with the youth's Ideal Self Goal and "+ 3" indicates behaviors

[5] This system is used to collect data on and analyze HRV in a non-invasive manner while estimating the function of the sympathetic and parasympathetic nervous systems. The system fulfills the requirements of the Heart Rate Variability Standards (Task Force of the European Society of Cardiology and Task Force of the European Society of Cardiology, 1996).

[6] RMSSD is a stable (Li et al., 2009) and valid (Thayer & Sternberg, 2010) time-domain measure of HRV.

that likely result in an impression by others that is congruent with the youth's Ideal Self Goal. "0" indicates a neutral rating and reflects "no evidence either way") (see Appendix). Self- and clinician ratings of Ideal Self Goal mastery was collected from youth in the first treatment group while they were receiving the treatment and from youth in the second treatment group while they were receiving the treatment.

Distal Outcome Measures

ISG-CM was designed to primarily improve ED and social functioning. To determine treatment effects on these outcomes, data on distal measures of ED and social functioning was collected from both participating youth and adults working with the youth, that is PMCJF staff and teachers. To determine treatment effects on outcomes related to ED and social functioning, data on distal measures of treatment effects were also collected on aggression, ADHD and ODD symptoms, academic functioning, and externalizing/internalizing symptoms. Data on secondary outcomes was collected from both participating youth and PMCJF staff and teachers.

Difficulties in Emotion Regulation Scale

The 36-item Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004, see Appendix) is a measure of self-rated ED. The DERS subscales are Lack of Emotional Awareness (Awareness), Lack of Emotional Clarity (Clarity), Difficulties Engaging in Goal-Directed Behavior (Goals), Impulse Control Difficulties (Impulse), Nonacceptance of Emotional Responses (Nonacceptance), and Limited Access to Emotion Regulation Strategies (Strategies). Data on construct and predictive validity as well as on internal consistency and test–retest reliability (Adrian et al., 2009; Bunford et

al., 2014; Bunford et al., 2015; Vasilev et al., 2009; Weinberg & Klonsky, 2009) indicate that the DERS has adequate psychometric properties. Higher values indicate more ED. Data on the DERS was collected from all youth at T1 and T2, and from youth in the second treatment group at T3.

Emotion Regulation Index for Children and Adolescents

The 16-item Emotion Regulation Index for Children and Adolescents (ERICA; MacDermott et al., 2010, see Appendix) is a measure of self-rated emotion regulation. The ERICA subscales are Emotional Control (Emotional Control), Emotional Self-Awareness (Self-Awareness), and Situational Responsiveness (Situational Responsiveness). Data on construct and convergent validity as well as internal consistency and test–retest reliability (Bunford et al., 2014; Bunford et al., 2015; MacDermott et al., 2010) indicate that the ERICA has adequate psychometric properties. Higher values indicate less ED. Data on the ERICA was collected from all youth at T1 and T2, and from youth in the second treatment group at T3.

Social Skills Improvement System-RS – Self-Report

The Social Skills Improvement System-RS – Self-Report (SSIS-RS; Elliott & Gresham, 2008) is a measure of self-rated social skills (46 items) and problem behaviors (29 items). Data on internal consistency and test-retest reliability (Gresham, Elliott, Vance, & Cook, 2011) indicate that the self-rated subscales have adequate psychometric properties. Higher values indicate better social skills and more problem behaviors. Data on the SSIS-RS was collected from all youth at T1 and T2, and from youth in the second treatment group at T3.

Buss–Perry Aggression Questionnaire

The 29-item *Buss–Perry Aggression Questionnaire* (BPAQ; Buss & Perry, 1992, see Appendix) is a measure of self-rated anger, hostility, physical aggression, and verbal aggression. Data on internal consistency and test-retest reliability (Buss & Perry, 1992) indicate that the BPAQ has adequate psychometric properties. Higher values indicate more aggression. Data on the BPAQ was collected from all youth at T1 and T2, and from youth in the second treatment group at T3.

Modified Overt Aggression Scale

The 25-item Modified Overt Aggression Scale (MOAS; Coccaro, Harvey, Kupsaw-Lawrence, Herbert, & Bernstein, 1991, see Appendix) is a measure of self- and staff-rated severity, type, and frequency of verbal aggression and physical aggression against self, objects, and others. Data on inter-rater reliability across clinicians and test–retest reliability (Suris et al., 2005) indicate that the MOAS has adequate psychometric properties. Higher values indicate more aggression. Data on the MOAS (both self- and staff-report) was collected from all youth at T1 and T2, and from youth in the second treatment group at T3.

PMCJF Behavior Points

Youth behavior and participation in PMCJF activities is evaluated on a point-system as part of regular facility functions. The structure of the point system is as follows: All youth begin their day having 20 points and can lose points for misbehavior or insufficient participation in facility activities. During the first half of each day, youth have three points for school, two points for daily living, two points for following facility

rules and three points for not having any incident reports (incident reports are written by facility staff in the case of serious aggression or rule-violations). During the second half of the day, youth have three points for following rules in groups, two points for daily living, two points for following facility rules and three points for not having any incident reports. Should a youth not lose any points during either the first or the second half of the day, he retains all 20 points for that day. Youth who retain all 20 points are considered to have "made their day" which, in turn, makes them eligible for two additional points, totaling to 22 points (youth can either exchange some of their 20 points for privileges or items from the commissary – in which case they do not earn two extra points, or they can keep their 20 points – in which case they earn two extra points). In addition to being able to exchange points for privileges or items from the commissary, youths' behavior points are used as an index of youths' compliance with, participation in, and progression with the program that is then reported to their probation officers and ultimately the courts and thus taken into account for decisions about youths' release. PMCJF behavior points were available for *all* youth for the duration of the first treatment group and were available for youth assigned to the second treatment group for the duration of the second treatment group.

Social and Mood Ratings

The six-item Social and Mood Ratings (SMR; Bunford & Evans, unpublished, see Appendix) is a measure of staff- and teacher-rated ED and social impairment. The SMR was created for this study. Higher values on both the emotion regulation and the social functioning items indicate better emotion regulation and better social functioning,

respectively. Data on the SMR (both staff- and teacher-report) was collected for all youth at T1 and T2, and for youth in the second treatment group at T3. Staff ratings on the SMR were also collected twice a day, two times a week for youth in the first treatment group while they were receiving the treatment and from youth in the second treatment group while they were receiving the treatment. Staff ratings on the SMR during treatment were not collected for youth in the control group so as to minimize burden on PMCJF staff.

Disruptive Behavior Disorders Rating Scale – Teacher Report

See Disruptive Behavior Disorders Rating Scale – Parent Report (DBD; Pelham et al., 1992, see Appendix) above for content. Data on the convergent validity (Molina, Pelham, Blumenthal, & Galiszevki, 1998) and reliability (Evans, et al., 2013) indicate that the inattention (IA), hyperactivity/impulsivity (H/I), and ODD subscales of the DBD–Teacher have adequate psychometric properties. Only the ADHD and the ODD items were included in this study. Higher values indicate more frequent and severe symptoms. Data on the DBD–Teacher was collected for all youth at T1 and T2, and for youth in the second treatment group at T3.

Impairment Rating Scale

The Impairment Rating Scale (IRS; Fabiano et al., 2006, see Appendix) is a measure of teacher-rated impairment in seven domains of functioning (e.g., relationships with peers, siblings, and adults; family functioning; academic progress; self-esteem) and global impairment. Data on convergent validity and internal consistency (Fabiano et al., 2006) indicate that the IRS has adequate psychometric properties. Higher values indicate

more impairment/a greater need for special services or treatment. Data on the IRS was collected for all youth at T1 and T2, and for youth in the second treatment group at T3.

Satisfaction and Preferred Treatment Elements

Two of the primary aims in the current study were to determine if youth found ISG-CM satisfactory and which treatment elements they liked and which ones they found beneficial/helpful. To address these aims, youth in the first treatment group completed ratings of treatment satisfaction and preferred treatment elements at T2 and youth in the second treatment group completed the same ratings at T3.

Satisfaction Questionnaire

A nine-item Satisfaction Questionnaire (Bunford & Evans, unpublished, see Appendix) was used to measure treatment satisfaction. The Satisfaction Questionnaire was adapted from Nguyen et al (1983) and created for this study. Higher values indicate better satisfaction.

Preferred Treatment Elements Questionnaire

The Preferred Treatment Elements Questionnaire (PTEQ; Bunford & Evans, unpublished, see Appendix) was used to measure youth perception of treatment elements with regard to the degree to which those are likeable and beneficial/helpful. The PTEQ was created for this study.

CHAPTER 3: RESULTS

Youth-Rated Satisfaction with ISG-CM

Descriptive analyses were conducted to examine youth ratings on the Satisfaction Questionnaire. Results indicated that, on average, youth found ISG-CM satisfactory (M = 31.58, SD = 4.76). With the highest possible rating being 36, one youth rated ISG-CM as 21, one as 25, and the remaining ten as ≥28 (Table 7). Upon further examining the data, among the ten youth who rated ISG-CM as ≥28, all rated each item as ≥3 on the four-point scale. Although the anchors for ratings of ≥3 are different across items, some examples of ratings of "3" indicate youth judged the quality of the services they received as good, reported that they generally got the kind of services they wanted, thought that the program met most of their needs, and felt mostly satisfied with the amount of help they received. The only item that was rated less than "3" by more than one youth was the item "If you were to seek help again, would you come back to our program?". In fact, other than this item and the item "In an overall, general sense, how satisfied are you with the service you have received", all 12 youth rated all items as ≥3 with the exception of one youth who rated six out of nine items as "2". Thus, there did not appear to be any specific domains of satisfaction that were rated lower than the rest. Conversely, all but one youth rated the item "How would you rate the quality of service you have received?" as "4" or "excellent", indicating that this domain of satisfaction was rated generally higher than the rest (the item that was second most frequently rated as "4" was rated as such by nine youth).

Youth-Rated Liked and Beneficial/Helpful Elements of ISG-CM

The number of youth who rated each treatment element as first, second, and third with regard to likeability and with regard to it being beneficial/helpful was counted. Psychoeducation was rated as the most liked element by one youth, Phase II group sessions were rated as the most liked by eight youth, and out-of-group mindfulness mediation was rated as the most liked by one youth (two youth did not provide ratings). As some examples of youths' explanation of their ratings, they noted that they liked the Psychoeducation element because "I liked setting the goals and helping me learn what I need to do [to achieve those]." Youth noted they liked Phase II group sessions because "It was a very fun way to learn how to get along with other kids, because you will get aggravated yet have to remember your goals." Another youth noted: "In [treatment team's] practice activities we worked hard and positively on our goals we set. [Treatment team] showed us how to work on our goals in every way. We got to rate our self then worked with our peers and counselors. I learned a lot." Finally, "I feel like these helped me get along better with others without getting mad." Regarding mindfulness mediation, youth noted that they liked it because "After [treatment team] showed us how to use the mindfulness mediation I put it to use every day, I used it and practiced it. It helps me every day. Outside of the facility, I will still be able to use it whenever I need it." Another youth stated "It would help me calm down when I have been angry." Similarly, two other youth reported that "It helps me calm down when I am mad or feel like doing something stupid. It also helps me go to sleep at night" and that "I like this because it calms me down."

Psychoeducation was rated as the most beneficial/helpful element by six youth, Phase II group sessions were rated as the most beneficial/helpful by five youth, and out-of-group mindfulness mediation was rated as the most beneficial/helpful by one youth (three youth did not provide ratings and the rest provided ratings that indicated they found two of the three elements equally beneficial/helpful) (Table 7). As some examples of youths' explanation of their ratings, youth noted they found Phase II group sessions beneficial/helpful because "I met my goals each group and learned how to control my emotions better." Another youth stated "It was helpful because we got feedback based on our goals." Regarding mindfulness meditation, youth reported "The mindfulness meditation has helped me with my anger. It helps me every day. I sleep with it. I feel like it will continue to help me even after group".

One question on the PTEQ is whether there was anything about the treatment that the questionnaire does not ask respondents about but that they really liked. In response to this question, one youth noted that "You were all very fair and respectful and insightful about others' needs and wants." Another youth stated "[Treatment team] positively helped me with my goals. My goals, I will be able to use every day inside and outside of the facility."

ISG-CM Effects on Proximal Measures of ED and Social Impairment

Proximal measures of ED and social impairment were HRV collected prior to, during, and following the frustration task and weekly self- and clinician-ratings of Ideal Self Goal mastery. Data on HRV were available only pre- and post-treatment but not for youth in the waitlist group (i.e., HRV data was collected at T1 and T2 for youth in the

first treatment group and at T2 and T3 for youth in the second treatment group. Due to attrition and unusable data, this resulted in a total $n = 6$ for these analyses). Data on Ideal Self Goals were available for youth while in ISG-CM.

ISG-CM Effects on Proximal Measures of ED

To examine changes in HRV over time, pre-post changes in mean HRV were examined for all three components of the experimental task including the baseline, frustration task, and recovery periods. Given small sample size and no HRV data for youth while not in their respective treatment groups, repeated measures effect sizes were calculated, in line with Gibbons, Hedeker, and Davis (1993) using the following formula: $((M_{\text{post, baseline}} - M_{\text{pre, baseline}})/SD_{\text{baseline}}))$; $((M_{\text{post, frustration}} - M_{\text{pre, frustration}})/SD_{\text{frustration}}))$; $((M_{\text{post, recovery}} - M_{\text{pre, recovery}})/SD_{\text{recovery}}))$. Here, the numerator is the sample mean change, or the mean difference between pre- and post-treatment scores, and the denominator is the sample standard deviation of change scores. In addition, to examine frustratibility, recovery, and return to baseline, difference scores were calculated to index the relative differences between baseline and frustration, frustration and recovery, and baseline and recovery elements, respectively. Then, repeated measures effect sizes were calculated in line with Gibbons, Hedeker, and Davis (1993) using the following formula = $((M_{\text{post, frustratibility}} - M_{\text{pre, frustratibility}})/SD_{\text{frustratibility}}))$; $((M_{\text{post, recovery}} - M_{\text{pre, recovery}})/SD_{\text{recovery}}))$; $((M_{\text{post, return to baseline}} - M_{\text{pre, return to baseline}})/SD_{\text{pre}}))$.

Comparing HRV at baseline, frustration task and recovery periods, all pre-treatment measurements were lower (or comparable, as in the case of pre-treatment baseline relative to post-treatment frustration and recovery relative) than their post-

treatment counterparts (indicating greater ED at pre- relative to post-treatment), with the largest effect size corresponding to differences in baseline HRV (d = .80; large), followed by differences in frustration task HRV (d = .38; medium), and by differences in recovery HRV (d = .22; small) (Figure 2).

Effect sizes were also calculated to compare pre- to post-treatment difference scores pertaining to frustratibility, recovery, and return to baseline. Difference scores pertaining to frustratibility were calculated to index the degree to which youth exhibited decreased HRV (indicating greater ED) during frustration relative to baseline. The magnitude of the difference in frustratibility from pre- to post-treatment was medium (d = .49). Difference scores pertaining to recovery were calculated to index the degree to which youth exhibited decreased HRV (indicating greater ED) during recovery relative to frustration. The magnitude of the pre- to post-treatment difference in recovery was negligible (d = .01). Finally, difference scores pertaining to return to baseline were calculated to index the degree to which youth exhibited decreased HRV (indicating greater ED) during recovery relative to baseline. The magnitude of the difference in return to baseline from pre- to post-treatment was medium (d = .42).

ISG-CM Effects on Proximal Measures of Social Impairment

Youth and clinicians rated youth behavior with regard to Ideal Self Goals twice per group session (i.e., first feedback and second feedback), during two group sessions per week. Two types of variables were derived from these ratings and those were self- and clinician ratings of Ideal Self Goal mastery and the relative similarity/difference between self- and clinician ratings. To evaluate changes during treatment in the former

type of variable, that is in the degree to which youth exhibited improvements in their ability to behave in a manner consistent with their Goals, changes in self- and clinician ratings were visually examined for trends over time (i.e., changes across all group sessions). To evaluate the degree to which youth exhibited improvements in their ability to accurately assess their behavior, difference scores were calculated by subtracting self- from clinician ratings. Then, changes in difference scores were examined visually for trends over time (i.e., changes across all group sessions). For both variables (i.e., changes in behavior and changes in accuracy), self- and clinician-ratings were averaged across youth in the treatment condition for the first feedback and across youth for the second feedback, for each group session. Data was analyzed separately for the first and second feedbacks as the first feedback reflects a conceptually different construct relative to the second feedback. At the time of the first feedback, youth have not been previously coached or given feedback since the preceding group session. However, at the time of the second feedback, youth were coached or given feedback only 15 to 20 minutes ago. Given this conceptual difference, it would stand to reason that self- and clinician ratings are lower for the first feedbacks relative to the second feedbacks. Indeed, paired samples *t*-tests comparing both average self- and clinician ratings across the first feedbacks to the second feedbacks indicated that both youth and clinicians rated youth behavior as better during the second feedback (self-ratings: $t[12] = -4.56$, $p = .001$; clinician ratings: $t[12] = -2.63$, $p = .022$).

With regard to youth behaving in a manner that is consistent with their Goals, no changes were observed in self-ratings of behavior for the first or the second feedback

across group sessions. No changes were observed in clinician-ratings of behavior for the second feedback across group sessions. Clinician-ratings of behavior improved, however for the first feedback across group sessions. Although a period of deterioration occurred during group sessions eight and nine, ratings continued to increase after group session nine (Figure 3).

Regarding changes over time in the accuracy of youths' self-ratings for the *first feedback*, youth rated their behavior more positively than clinicians until the fifth group session. After the fifth session, youth rated their behavior increasingly more negatively than clinicians. Interestingly for the *second feedback*, youth never rated their behavior more positively than clinicians. Youth initially rated their behavior similarly, with a trend starting with the fifth group session indicating they rated their behavior increasingly more negatively (Figure 4).

ISG-CM Effects on Distal Measures of ED and Social Impairment

Distal measures of ED were self-ratings on the ERICA, DERS, and BPAQ anger subscale, staff-ratings of ED on the SMR, and teacher ratings of ED on the SMR. Distal measures of social impairment were self-ratings on the SSIS-RS, staff-ratings of social impairment on the SMR, and teacher ratings of social impairment on the SMR. Mixed between-within and within subjects analyses were conducted on these indices to evaluate changes in primary treatment targets, ED and social impairment.

Data on additional indices of ED and social impairment were staff-ratings on the SMR and PMCJF behavior points collected *during* treatment. Staff-ratings of ED and social impairment *during* treatment were available for youth in ISG-CM only (as noted

above, to minimize burden on facility staff) and behavior points were available for both ISG-CM and control groups. To analyze these data, trends in staff-rated ED and social impairment on the SMR as well as in PMCJF behavior points during the treatment (for ISG-CM only) were examined.

To test for differences between ISG-CM and waitlist control on ED and social impairment, mixed between-within subjects analyses of variance (ANOVA) were conducted. The models included group assignment (ISG-CM vs. waitlist control) as the between-subjects factor, time (pre- to post-treatment) as the within-subjects factor and the interaction between group and time as an indication of treatment response. Group x Time interactions were tested using pre-treatment and post-treatment data for six youth who were assigned to ISG-CM and for six youth who were assigned to waitlist. The Group x Time interaction was the main outcome of interest as a significant interaction is indicative of groups changing differently over time. Given small sample size, both statistical significance and effect size magnitude were considered for determining when to conduct follow-up analyses. The Group x Time interactions were considered *meaningful if* $p \leq .1$ (given that $p \leq .1$ is often used for determining marginal significance) and/or $\eta_p^2 \geq .13$ (given that $\eta_p^2 \geq \sim.14$ is used as a cutoff for a large effect [Watson, 2016] and that in light of small sample, the ~.14 was interpreted liberally as including .13). Because η_p^2 corresponds to the portion of the variance in the dependent variable that is explained by group membership (treatment vs. control) but does not allow for estimating the magnitude of the difference between groups, meaningful Group x Time interactions were followed with calculations of two types of effect size. The first (Cohen's *D*),

comparing the treatment group pre-treatment means to the control group pre-treatment means was calculated, in line with Hedges (1981; 1982) using the following formula: $(M_{\text{pre, treatment group}} - M_{\text{pre, waitlist group}})/(\text{SQRT}((SD_{\text{pre, treatment group}}*SD_{\text{pre, treatment group}}+SD_{\text{pre, waitlist group}}*SD_{\text{pre, waitlist group}})/2))$ and comparing the treatment group post-treatment means to the control group post-treatment means was calculated using: $(M_{\text{post, treatment group}} - M_{\text{post, waitlist group}})/(\text{SQRT}((SD_{\text{post, treatment group}}*SD_{\text{post, treatment group}}+SD_{\text{post, waitlist group}}*SD_{\text{post, waitlist group}})/2))$. The second, or repeated measures effect size, was calculated in line with Becker (1988) using the following formula: $((M_{\text{post, treatment group}} - M_{\text{pre, treatment group}})/SD_{\text{pre, treatment group}})) - ((M_{\text{post, waitlist group}} - M_{\text{pre, waitlist group}})/SD_{\text{pre, waitlist group}}))$. Group x Time interaction outcomes are presented in Table 8 with corresponding descriptive statistics and effect sizes presented in Table 9.

To test for differences across baseline (T1), pre-treatment (T2), and post-treatment (T3) using participants as their own controls, within subjects ANOVAs were conducted. Similar to Group x Time interactions, statistical significance and effect size magnitude were considered ($p \leq .1$ and/or $\eta_p^2 \geq .13$) for determining when to conduct follow-up analyses, which consisted in calculations of repeated measures effect size in line with Gibbons, Hedeker, and Davis (1993) using the following formula: $(M_{T2} - M_{T1})/SD_{T2,\ T1})$ and $(M_{T3} - M_{T2})/SD_{T3,\ T2})$. Here, the numerator is the sample mean change, or the mean difference between T1 and T2 (and T2 and T3) scores, and the denominator is the sample standard deviation of change scores. Effect sizes and descriptive statistics corresponding to meaningful within-subjects comparisons are presented in Table 10.

ISG-CM Effects on Distal Measures of ED

Distal measures of ED were self-ratings on the ERICA, DERS, and BPAQ anger subscale, staff-ratings on the SMR, and teacher ratings on the SMR collected from both groups at all time-points. The anger subscale is conceptualized as an index of ED given its content (e.g., "I flare up quickly but get over it quickly", "When frustrated, I let my irritation show", "I sometimes feel like a powder keg ready to explode", "I am an even-tempered person", etc.). Additional distal measures of ED were staff-ratings on the SMR collected during treatment from youth in ISG-CM only. Below, only changes considered meaningful on mixed between-within and within analyses are reported. Results that did not reach the stated threshold ($p \leq .1$ and/or $\eta_p^2 \geq .13$) are not reported. Mixed between-within and within subjects analyses are followed by a discussion of trends during treatment on staff-rated SMR scores.

Regarding differential changes between groups, across the ratings of two teachers, the Group x Time interaction for teacher-rated emotion regulation on the SMR was meaningful based on effect size estimates ($p = .145$, $\eta_p^2 = .18$ and $p = .180$, $\eta_p^2 = .16$). Cohen's *D* calculations revealed a small-medium difference between groups pre-treatment and a large difference between groups post-treatment for the ratings of both teachers and with the treatment group exhibiting better emotion regulation at both time-points (Figure 5). The repeated measures effect sizes corresponding to these differences were medium for one teacher ($d = .31$) and small for the other teacher ($d = .22$) (Table 9).

Regarding changes across three time-points, there was a meaningful difference based on effect size estimates in self-rated Impulse on the DERS, with youth exhibiting

less ED over time (p = .347, η_p^2 = .30). The repeated measures effect sizes were small from T1 to T2 (d = -.25) but negligible from T2 to T3 (d = -.06). Time explained 30% of the variance in change in behavioral dyscontrol in the face of strong emotions (Figure 6). There was also a meaningful difference based on effect size estimates in self-rated anger on the BPAQ, with youth reporting the same amount of anger from T1 to T2 but less from T2 to T3 (p = .454, η_p^2 = .18). The repeated measures effect sizes were negligible from T1 to T2 (d = -.02) but large from T2 to T3 (d = -1.16). Time explained 18% of the variance in changes in anger (Figure 6).

In addition to the between-within and the within subjects analyses, trends in staff-rated emotion regulation on the SMR during the treatment (for ISG-CM only) were examined. Of interest was change over time. Ratings for these analyses were averaged across youth and across two days/week with ratings from morning staff and afternoon staff analyzed separately. Results indicated no notable trends with regard to morning or afternoon staff-rated ED (Figure 7). Interestingly, however, a relative drop in staff-rated emotion regulation was observed during the fourth week (day two for morning shift ratings). The timing of this deterioration coincided with a period of deterioration in clinician-ratings of Ideal Self Goal mastery (during group sessions eight and nine, see above and *Discussion* on why these findings may be noteworthy).

ISG-CM Effects on Distal Measures of Social Impairment

Distal measures of social impairment were self-ratings on the SSIS-RS, staff-ratings on the SMR, and teacher ratings on the SMR collected from both groups at all time-points. Additional distal measures of social impairment were staff-ratings on the

SMR collected during treatment from youth in ISG-CM only and PMCJF behavior points collected during treatment from both groups. Below, only changes considered meaningful on mixed between-within and within analyses are reported. Results that did not reach the stated threshold ($p \leq .1$ and/or $\eta_p^2 \geq .13$) are not reported. Mixed between-within and within analyses are followed by a discussion of trends during treatment on staff-rated SMR scores and PMCJF behavior points.

Regarding differential changes between groups, the Group x Time interaction for staff- ($p = .205$, $\eta_p^2 = .17$) and teacher-rated ($p = .108$, $\eta_p^2 = .22$) social functioning on the SMR were meaningful based on effect size estimates. For staff ratings, Cohen's *D* calculations revealed a small difference between groups pre-treatment and a large difference between groups post-treatment, with the treatment group exhibiting better social functioning at both time-points (Figure 5). For teacher ratings, Cohen's *D* calculations revealed a small difference between groups pre-treatment and a medium-large difference between groups post-treatment, with the treatment group exhibiting worse social functioning at pre-treatment and better social functioning at post-treatment (Figure 5). The repeated measures effect sizes for staff and teacher ratings were large ($d = .80$ and .84) (Table 9).

Regarding changes across three time-points, there was a meaningful difference based on effect size estimates in teacher-rated overall impairment on the IRS, with youth exhibiting more impairment from T1 to T2 and less impairment from T2 to T3 ($p = .460$, $\eta_p^2 = .19$). The repeated measures effect size was small for T1 to T2 ($d = .28$) and medium from T2 to T3 ($d = -.45$) (Table 10). Time explained 19% of the variance in

change in overall impairment (Figure 6). There was also a meaningful difference based on effect size estimates in teacher-rated teacher-child relationship impairment on the IRS and on teacher-rated classroom impairment on the IRS. On both indices, youth exhibited more impairment from T1 to T2 and less impairment from T2 to T3 (*p*s = .332, η_p^2 = .89; 574, η_p^2 = .17, respectively). The repeated measures effect size corresponding to teacher-child relationship impairment was large for T1 to T2 (*d* = .92) and small from T2 to T3 (*d* = -.25) (Table 10). The repeated measures effect size corresponding to classroom impairment was negligible for T1 to T2 (*d* = .10) and large from T2 to T3 (*d* = -.71) (Table 10). Time explained 89 and 17% of the variance in change in teacher-rated teacher-child relationship impairment and in classroom impairment, respectively (Figure 6).

In addition to the between-within and the within subjects analyses, changes in staff-rated social functioning on the SMR during the treatment (for ISG-CM only) were also visually examined for trends. Of interest was change over time. Similar to staff-rated ED, ratings were averaged across youth and across two days/week with ratings from morning staff and afternoon staff analyzed separately. Similar to ED, no notable trends were observed with regard to morning or afternoon staff-rated social functioning on the SMR (Figure 7). Also similar to ED, a relative drop in staff-rated social functioning was observed during the fourth week (day two for morning and afternoon shift ratings). The timing of this deterioration coincided with a period of deterioration in clinician-ratings of Ideal Self Goal mastery (see above and *Discussion* on why these findings may be noteworthy).

Changes in PMCJF behavior points during the treatment (for both ISG-CM and control groups) were visually examined for trends. Of interest was change over time. Data was analyzed in two ways: first, the behavior points of the ISG-CM group were compared to the waitlist group. For these comparisons, points were averaged across youth and across seven days/week. Second, youths' points while they were waitlisted were compared to their points while they were in ISG-CM. For these analyses, points were considered in terms of the relative portion of successful and unsuccessful days (defined as the youth earning ≤14 points as for youth to lose six points, they would have had to receive zero points for one entire activity [e.g., school] in the morning and one in the afternoon [e.g., groups]).

With regard to behavior points comparing ISG-CM to control, youth in the waitlist group earned more points during the first week, youth in the waitlist and treatment groups earned the same amount of points during the second week, and youth in ISG-CM consistently earned more points during the rest of the weeks (see Figure 8). With regard to behavior points comparing youths' points while they were waitlisted to their points while they were in ISG-CM, encouraging trends were observable for the relative portion of unsuccessful days that youth had before compared to during the treatment (Figure 8). Three of the five youth had a smaller portion of unsuccessful days during the treatment than before; for one of these three youth, the portion was almost halved and for another it was reduced by 4/5ths. Although two youth had a larger portion of unsuccessful days during the treatment than before, it is important to note that both had zero unsuccessful days before treatment. Therefore, it was not possible for these youth to

have less unsuccessful days. In addition, the degree to which these two youths' unsuccessful days increased was lower than the portion of unsuccessful days experienced by the other youth before the treatment.

ISG-CM Effects on Aggression, ADHD and ODD symptoms, Academic Impairment, and Externalizing/Internalizing Problems

Distal measures of aggression were self-ratings on the BPAQ and staff-ratings on the MOAS; of ADHD and ODD symptoms were teacher ratings on the DBD; of academic impairment were teacher ratings on the IRS; and of externalizing/internalizing problems were self-ratings on the SSIS-RS. Below, only changes considered meaningful on mixed between-within and within analyses (conducted in the same way as for ED and social impairment) are reported. Results that did not reach the stated threshold ($p \leq .1$ and/or $\eta_p^2 \geq .13$) are not reported.

ISG-CM Effects on Aggression

Regarding differential changes between groups, the Group x Time interaction for self- and staff-rated verbal aggression on the BPAQ ($p = .263$, $\eta_p^2 = .18$) and the MOAS ($p = .119$, $\eta_p^2 = .23$), respectively, as well as for staff-rated aggression against property on the MOAS ($p = .173$, $\eta_p^2 = .18$) were meaningful based on effect size estimates. For both self- and staff-rated verbal aggression, Cohen's *D* calculations revealed a small-medium difference between groups pre-treatment and a large difference between groups post-treatment. The treatment group exhibited less self-rated aggression at both time-points (Figure 9) and the treatment group exhibited more staff-rated aggression at pre-treatment and less staff-rated aggression at post-treatment (Figure 9). For staff-rated aggression

against property, there was a large difference between groups pre-treatment and a medium-large difference between groups post-treatment (Cohen's *D*). The treatment group exhibited more aggression pre-treatment and less aggression post-treatment (Figure 9). The repeated measures effect sizes for self-rated verbal aggression was large (d = .80), for staff-rated verbal aggression was small (d = .23), and for staff-rated aggression against property was medium (d = .44) (Table 9).

Regarding changes across three time-points, there was a meaningful difference based on effect size estimates in self-rated verbal aggression on the BPAQ (p = .554, η_p^2 = .18), with youth exhibiting more aggression from T1 to T2 and less aggression from T2 to T3. The repeated measures effect size corresponding to self-rated verbal aggression was small for T1 to T2 (d = .26) and large from T2 to T3 (d = -1.64) (Table 10). Time explained 18% of the variance in change in self-rated verbal aggression (Figure 6). This finding was replicated based on effect size estimates with staff-rated verbal aggression on the MOAS (p = .359, η_p^2 = .23), with youth exhibiting more verbal aggression from T1 to T2 but a little less aggression from T2 to T3. The repeated measures effect size corresponding to staff-rated verbal aggression was large for T1 to T2 (d = .56) and negligible from T2 to T3 (d = -.11) (Table 10). Time explained 23% of the variance in change in staff-rated verbal aggression (Figure 6). Similarly, there was a meaningful difference based on effect size estimates in staff-rated aggression against property on the MOAS (p = .570, η_p^2 = .13). Although youth exhibited more aggression from T1 to T2 and from T2 to T3, the increase from T1 to T2 was greater than from T2 to T3 (Cohen's *D*s -0.61 and -0.15). The repeated measures effect size corresponding to staff-rated

aggression against property was medium for T1 to T2 (d = .45) and negligible from T2 to T3 (d = .11) (Table 10). Time explained 13% of the variance in change in aggression against property (Figure 6).

ISG-CM Effects on ADHD and ODD Symptoms

Regarding differential changes between groups, the Group x Time interaction for teacher-rated IA on the DBD was statistically significant (p = .030, η_p^2 = .51). There was a small difference between groups pre-treatment and a medium difference post-treatment (Cohen's D), with the treatment group exhibiting less IA at both time-points (Figure 10). The repeated measures effect size for IA was large (d = .65) (Table 9). The Group x Time interaction for teacher-rated H/I on the DBD was also meaningful based on effect size estimates (p = .460, η_p^2 = .14). There was a small pre-treatment between-group difference and a large post-treatment difference, with the treatment group exhibiting more H/I at pre-treatment but less at post-treatment (Figure 10). The repeated measures effect size for H/I was small (d = .20) (Table 9).

ISG-CM Effects on Academic Impairment

The Group x Time interaction for teacher-rated academic functioning on the IRS was meaningful based on effect size estimates (p = .261, η_p^2 = .14). There was a small-medium pre-treatment between-group difference but a medium-large post-treatment difference (Cohen's D), with the treatment group exhibiting more impairment at pre-treatment but less at post-treatment (Figure 11). The repeated measures effect size for academic impairment was large (d = 2.77) (Table 9).

ISG-CM Effects on Externalizing/Internalizing Problems

The Group x Time interaction for self-rated internalizing problems on the SSIS was meaningful based on effect size estimates ($p = .223$, $\eta_p^2 = .14$). There was a medium pre-treatment between-group difference and a small post-treatment difference (Cohen's *D*), with the treatment group having more internalizing problems at pre-treatment but less at post-treatment (Figure 12). The repeated measures effect size for internalizing problems was large ($d = 2.27$) (Table 9).

CHAPTER 4: DISCUSSION

The primary goals of the current study were to examine two primary research questions about ISG-CM, delivered in a juvenile facility center and using multi-method and multi-informant measurement. The first outcome of interest was participant satisfaction with ISG-CM and the second was the preliminary effectiveness of ISG-CM in improving ED, social impairment, and functioning in related domains. Although the current pilot clinical trial was conducted with a small sample, the obtained results provide support for participant satisfaction with ISG-CM. The results also provide preliminary support for the beneficial effectiveness of ISG-CM on improving accuracy of self-appraisal of behavior, aggression, and daily behavior at the facility. These findings have implications for further treatment development and evaluation work to meet the needs of detained juvenile offenders with externalizing disorders.

Satisfaction and Preferred ISG-CM Elements

The findings of the current study indicate that youth were highly satisfied with ISG-CM and that all components of the treatment were either liked by youth, found to be beneficial by youth, or both. Regarding satisfaction, the majority of youth rated each measured aspect of satisfaction as at least “3” on a scale ranging from “1” to “4”, with ratings of “1” generally indicating that youth judged the quality of the services they received as poor and that the program met none of their needs and with ratings of “4” generally indicating that youth judged the quality of the services they received as excellent and that the program met almost all of their needs (the anchors are different across items). One youth’s satisfaction ratings of ISG-CM were as high as possible and

seven youths' ratings were almost as high as possible. The only item that was rated lower than "3" by more than one youth was, "If you were to seek help again, would you come back to our program?", and this may reflect that youth interpreted this item as involving returning to the detention facility, which seems an undesirable option.

Although youth did not like Psychoeducation, they did find it beneficial. The majority of youth liked the Phase II group sessions the most and more than half found these sessions to be the most beneficial. Although only one youth liked mindfulness meditation most and found it to be most beneficial, approximately half of the youth found this treatment element to be second most liked and to be second most beneficial. Youths' qualitative reports consistently indicated that they liked and found benefit/help in Phase II group activities primarily because of being able to get feedback on goals they have set for themselves and to practice/work towards such goals in a group or peer setting. Regarding the mindfulness meditation, as expected, youth consistently indicated that it helped them calm down and regulate their anger. Perhaps less expected was the finding that the qualitative reports of a large number of youth indicated that the mindfulness meditation element helped them with their sleep. This finding is not surprising, per se, given that it was expected that meditation practice would result in relaxation, which certainly can be helpful in improving sleep. What is surprising, however, is that many youth specifically and without prompting identified an improvement in sleep as a beneficial effect of the meditation element (during Psychoeducation, youth were informed that meditation may help with relaxation. They were not informed that it may help with sleep).

A measurement limitation was that youth did not have the option to indicate that they did not like a treatment element or that they did not find it beneficial/helpful. On the PTEQ, youth were asked to rank all three treatment elements but were not asked the dichotomous question of whether or not they liked the treatment element and of whether or not they found it beneficial/helpful. Thus, as an example, youth may have ranked an element as most liked not because they liked it but because they disliked it less than the other two elements. Of note however, youth had the opportunity to indicate which elements they did not like (this was only the case for Psychoeducation) or did not find beneficial/helpful (this was not the case for any of the elements) in their written comments.

Taken together, quantitative and qualitative findings related to satisfaction and preferred treatment elements indicate that *youth were satisfied with ISG-CM, and* either *liked or believed that they benefitted from* each of three *treatment elements*. In addition to reflecting positively on ISG-CM procedures related to goal setting, practice, and feedback, youths' reports also indicate that there may be utility in specifically linking mindfulness meditation practice to potential improvements in sleep in future iterations of ISG-CM.

Consistent and Feasible Delivery of ISG-CM

Although no formal measurement of treatment integrity was implemented, the treatment team was able to consistently implement ISG-CM in the current study. For both treatment groups, all sessions were held, all but one included both feedbacks with emotion ratings, and all were concluded with group mindfulness meditation. The

consistent support of the PMCJF director, who continuously communicated the importance of the present research to all staff, was central to such successful implementation of ISG-CM. One meaningful obstacle was youth being discharged from the facility while in treatment. Nevertheless, in total, *ISG-CM was consistently and feasibly delivered in a juvenile detention facility*.

Improvements in Accuracy of Self-Appraisal

The findings of the current study indicate that feedback and discussion of behavior in relation to social goals improve the accuracy of youths' self-appraisals. Specifically, although youth rated their behavior more positively than clinicians during the first feedbacks, *youth never rated their behavior more positively than clinicians during the second feedbacks*. These data are encouraging for several, related reasons. They at minimum indicate that youth were engaged in ISG-CM. This implication is significant as engagement with treatment is associated with benefitting from treatment (e.g., Fiorentine, Nakashima, & Anglin, 1999). Related, beyond mere engagement, these findings indicate that feedback and discussion of discrepancy between self- and clinician-ratings improves the accuracy of youths' self-appraisals. This implication is important as it reflects that the change that is intended to occur is indeed occurring between the two feedbacks and supports the notion behind ISG-CM as a TI. It stands to reason that over time, with sufficient practice and repetition, the observed increase in accuracy during the second feedbacks relative to the first feedbacks may attenuate as self-appraisal skills become automatized and improved. This is a testable hypothesis and the authors of future

studies may focus on identifying ways to ensure long-term effects of the feedbacks, perhaps with longer treatment sessions and with three feedbacks per group.

It is interesting that although no notable results were obtained with proximal indices of staff-rated ED and social functioning, there was a period of deterioration in clinician-rated Ideal Self Goal mastery and this period coincided with a period of deterioration in staff-rated ED and social impairment on the SMR. The former and latter ratings were completed independently and by different raters (clinicians vs. staff) and in different settings (during ISG-CM Phase II sessions vs. outside of ISG-CM Phase II sessions). Clinicians were unaware of staff ratings of ED and social functioning and staff were unaware of youths' clinician-rated performance in group. On one hand, the congruence between these measures may be due to chance. On the other hand, the congruence could reflect concurrent validity, especially insofar as these measures are sensitive to decrements in behaviors or functioning. In future studies investigators may focus on further testing the concurrent validity of clinician and staff ratings in these domains.

Of further note, the results obtained in the current study with regard to the deterioration youth exhibited while waitlisted are comparable to trends obtained by Schlichter and Horan (1981) in some of their data. Although improvement, stagnation, or deterioration is not discussed by the authors of any of the prior studies wherein treatments for youth in juvenile detention settings were evaluated, the data presented by Schlichter and Horan (1981) allow for concluding that their sample also exhibited decrements on indices of self-ratings of anger in response to tape-recorded provocative scenes, self-

ratings of the degree to which they desired to exhibit verbal aggression in response to tape-recorded provocative scenes, and self-ratings of pro-aggression cognitions when not in treatment (the data reported by Schlichter and Horan [1981] did not allow for estimating repeated measures within effect sizes so the magnitude of the decrements those authors observed is unclear).

Related, while at the PMCJF, clinical staff had frequent discussions with PMCJF staff. During these discussions, clinical staff learned about "life" at a detention facility, including that there are days that are a "bad day for an entire unit or the entire place" (e.g., when an extreme event happens) and manifest in youth collectively altering their behavior. It may have been that during the fourth week of ISG-CM, such an extreme event happened and both clinician-ratings of Ideal Self Goal mastery and staff-ratings on the SMR reflect collectively altered behavior. In future studies it may be helpful to measure the impact of context on youths' social functioning, such as disruptive events, changes in facility rules, staff, or teachers in order to better understand the outcomes that may be influenced by such events.

Improvements in Autonomic Rigidity

The current results indicate that youth may have exhibited *increased autonomic nervous system flexibility following treatment*. Specifically, pre- to post-treatment changes in HRV were observed, with youth exhibiting higher HRV post-treatment (indicating better emotion regulation) especially during baseline, as indicated by a large effect, and to a lesser extent during the frustration task and during recovery, as indicated by a medium and a small effect, respectively. Of import, all post-treatment values were

greater than – or comparable to – pre-treatment values indicating that even when frustrated, youth exhibited less autonomic arousal at post-treatment than when not frustrated at pre-treatment. Similarly, comparing the *relative* change in autonomic arousal across measurements, i.e., frustratibility, recovery, and return to baseline, pre- to post-treatment changes in frustratibility and return to baseline were observed, with *youth exhibiting less frustratibility and better return to baseline* post-treatment and as indicated by medium effects. Although these findings are encouraging, because the relevant measurements were not controlled, it would be premature to interpret these data to reflect on whether ISG-CM results in measurable changes in these psychophysiological indices of ED or on whether HRV as a psychophysiological index of ED is sensitive to the effects of ISG-CM. Related, accompanying the positive findings, it is of clinical concern that youth did not improve with regard to recovery, as indicated by a negligible effect. Additional research is needed to address the questions related to the relation between HRV as a psychophysiological index that is both changeable through ISG-CM and is a sensitive index of change as a result of ISG-CM. Additional research is also needed with an extended recovery period so as to improve the extent to which the experimental task is suitable to capture youths' complete return to baseline.

Improvements in Aggression, Anger, and Daily Behavior

Although not one of the primary outcomes of interest, *changes in aggression,* both self- and staff-rated, *are* perhaps the *most clinically meaningful*. Across analyses, participation in ISG-CM was associated with the absence of an increase in self- and staff-

rated verbal aggression and the absence of an increase in staff-rated aggression against property.

Youth in ISG-CM exhibited no changes in self-rated verbal aggression but youth in the waitlist group exhibited increases in verbal aggression, to a degree that corresponds to almost one point per item on the seven-point BPAQ (ranging from "extremely uncharacteristic of me" to "extremely characteristic of me"). The corresponding mixed between-within effect was large for self-rated verbal aggression, further underscoring that this change was clinically meaningful. Within-subjects results also indicated noteworthy changes in this domain following ISG-CM; while waitlisted youths' verbal aggression increased but then while in ISG-CM it decreased. The repeated measures effect corresponding to the increase while waitlisted was small and the one corresponding to the decrease while in ISG-CM was large, indicating that youths' verbal aggression first increased to an extent that may have been behaviorally and clinically noticeable but then decreased to a degree that was likely clinically meaningful. Importantly, these changes in self-rated verbal aggression were replicated with staff-rated verbal aggression with regard to the direction of the observed effects (but not the magnitude of the observed effects). Youth in ISG-CM exhibited a minimal decrease in staff-rated verbal aggression, whereas the average verbal aggression of youth in the waitlist group nearly tripled and the corresponding mixed between-within effect was small for staff-rated verbal aggression. The verbal aggression of youth in the waitlist group increased by over one units. Their pre-treatment verbal aggression ranged from "none" to "angry shouting, making personal insults and mild cursing" and their post-treatment verbal aggression ranged from "angry

shouting, making personal insults and mild cursing" to "severe insults, temper outbursts, and vicious cursing"). Thus, although a similar change was observed for staff-rated verbal aggression as for self-rated verbal aggression, the magnitude of change in the former was smaller, perhaps due to reputation bias. Specifically, staff may have associated some youth with the tendency to exhibit verbal aggression based on early their behavior and additional instances of non-aggressive communication would have been needed for such association to significantly change. Within-subjects results also indicated that youth exhibited an almost two unit increase in staff-rated verbal aggression while waitlisted (on average, at T1, their verbal aggression ranged from "none" to "angry shouting, making personal insults and mild cursing" whereas at T2 their verbal aggression ranged from "angry shouting, making personal insults and mild cursing" to "severe insults, temper outbursts, and vicious cursing"). Conversely, youth exhibited decreased staff-rated verbal aggression at T3. The repeated measures within effect corresponding to the increase while waitlisted was large and the one corresponding to the decrease while in ISG-CM was small. Thus, unlike self-rated verbal aggression, where youth first exhibited a small increase but then a large and likely clinically meaningful decrease, staff saw youth as exhibiting a behaviorally and clinically noticeable increase in aggression while not in treatment but then somewhat of an improvement while in treatment. The reputation bias discussed above may account for the relatively smaller magnitude of staff-rated change in response to treatment.

Changes in staff-rated aggression against property were similar to changes in staff-rated verbal aggression. Youth in ISG-CM exhibited decreased aggression against

property from pre- to post-treatment whereas youth in the waitlist group exhibited clinically meaningful increases in this domain (from “no aggression” to ranging from “ripping clothing, slamming doors, or urinating on the floor” to “defacing walls, kicking furniture, and throwing objects down”). The corresponding mixed between-within effect was medium for staff-rated aggression against property, further underscoring that this change was clinically meaningful. Although within-subjects results corresponding to staff-rated aggression against property indicated that all youth exhibited increased aggression over time, this increase was 1.5 units while waitlisted whereas it was less than 0.5 units while in ISG-CM. A one unit increase could indicate that youth increased from “not exhibiting aggression against property” to exhibiting aggression ranging from “ripping clothing, slamming doors, or urinating on the floor” to “defacing walls, kicking furniture, and throwing objects down”. Conversely, a 0.5 unit increase indicates that youth remained in the latter range. The repeated measures effect corresponding to the increase while waitlisted was medium and the one corresponding to the increase while in ISG-CM was small. Similar to staff-rated verbal aggression but unlike self-rated verbal aggression, staff saw youth as exhibiting a behaviorally and clinically noticeable increase in aggression against property while not in treatment but then somewhat of an improvement while in treatment, indicating consistency in staff-ratings across domains of aggression.

Another encouraging finding in a related domain was that while waitlisted, youths’ self-rated anger increased but while in treatment, it decreased to a clinically meaningful degree (on a seven-point scale ranging from “extremely uncharacteristic of

me" to "extremely characteristic of me", youths' improvement corresponded to a two point decrease in anger per item). The repeated measures effect corresponding to the increase while waitlisted was medium and the one corresponding to the decrease while in ISG-CM was large. This pattern is unlike staff-rated verbal aggression and aggression against property but similar to self-rated verbal aggression, indicating consistency in self-ratings across aggression and anger.

Finally, regarding daily behavior and participation at the facility, relative to the waitlist group, youth in ISG-CM first earned fewer points, then earned the same amount of points, and then earned consistently more points. Corresponding within subjects analyses also indicated that participation in ISG-CM was associated with a relative absence of unsuccessful days (in terms of behavior points).

Thus, across analyses, receipt of ISG-CM was associated with the absence of an increase in self- and staff-rated verbal aggression, in staff-rated aggression against property, and in anger as well as with improvement in daily behavior at the facility and participation in facility activities. Because behavior points were assigned by raters unaware of condition (i.e., PMCJF staff were unaware of group assignment), these data indicate that ISG-CM is associated with benefits (e.g., in institutional adjustment) that are observable outside of group sessions. The findings on aggression, anger, and daily behavior highlight the importance of these domains of functioning as potential behavioral and emotional targets for these youth. Thus, these domains of functioning may also be worthwhile to include as primary areas of interest in future iterations and evaluations of ISG-CM.

Of note, the results obtained in the current study are largely consistent with the encouraging findings obtained in some of the prior studies wherein treatments for youth in juvenile detention settings were evaluated. Specifically, those encouraging findings included reductions in aggression (Guerra & Slaby, 1990; Schlichter & Horan, 1981; Shelton et al., 2011) or in pro-aggression cognitions (Guerra & Slaby, 1990); improvements in daily behavior at the facility and participation in facility activities (Ross & McKay, 1976; Shelton et al., 2011; Trupin et al., 2002); and improvement in anger (Schlichter & Horan, 1981) or anger control (Shivrattan, 1988). As noted above, because effect size estimates were not reported in any of the prior studies and neither were *M* or *SD* values that would allow for the calculation of such estimates, it is difficult to compare the magnitude of the effects obtained in the current study to the magnitude of the effects obtained in prior studies. As also noted, Schlichter and Horan (1981) reported sufficient data to determine the magnitude of the effects they found. The majority of their effects were medium or large including those corresponding to improvement in the extent to which youth estimated they would feel angry in response to provocation, as well as in observer-rated verbal aggression during role-plays of provoking social interactions. Although some of the effects obtained in the current study are comparable to those obtained in this prior study with regard to magnitude, they were generally smaller. The effects obtained by Schlichter and Horan (1981) with regard to observer-rated aggression at the facility outside of treatment and experimental paradigms, were small-medium for verbal aggression and large for physical aggression. The effects obtained in the current study are comparable to those obtained by Schlichter and Horan (1981) in terms of these

indices of aggression. Schlichter and Horan (1981) obtained generally larger effects with regard to self-rated hypothetical anger and observer-rated aggression during experimental paradigms but obtained comparable effects with regard to observer-rated aggression during daily functioning at the facility. Arguably, these differences may be due to the inherent difference between estimated aggression to hypothetical scenarios and observed aggression during role-plays compared to actual aggression during actual social situations. Of import, the latter type of index, i.e., indices of actual aggression are more ecologically valid, not only further underscoring the meaningfulness of the current results but also highlighting their comparability to prior results on indices that are clinically meaningful.

Lack of Improvements in ED and Prosocial Behaviors

Youth did not exhibit changes in some of the behavioral and emotional targets of ISG-CM. Specifically, other than in the case of a few exceptions, youth did not exhibit improved emotion regulation and social functioning. Null findings were obtained with the DERS (except for Impulse, see above) and the ERICA as distal indices of emotion regulation and with the SSIS-RS as a distal index of social functioning. These data may reflect that ISG-CM does not result in measurable changes on these indices of ED and social impairment, indicating that there is a need for further modifying ISG-CM. Alternatively, these data may reflect that these measures are not sensitive to ISG-CM effects, indicating that in future studies of the treatment, they may not be useful to include as measures of primary outcomes. In support of the latter hypothesis, these results are consistent with findings obtained in prior evaluations of ISG (e.g., Evans et al., 2016)

wherein youth did not exhibit improved social functioning as indexed by the SSIS. These results are also consistent with the findings obtained in the prior studies wherein treatments for youth in juvenile detention settings were evaluated, with one exception being the study by Shivrattan (1988) who observed that social interactional training was associated with improvements in social skills on an observer-rated measure of conformity, considerateness, social self-control and unobtrusiveness. The negative findings obtained in the present study and prior studies (e.g., Evans et al., 2016), in combination with the positive findings obtained by Shivrattan (1988) may indicate that observational measures of social functioning may be more sensitive to treatment effects than rating scale measures. Thus, future evaluations of ISG-CM should include observational measures of social functioning to determine whether or not those are indeed sensitive to treatment effects.

As a final note regarding these measures being sensitive to treatment effects, it is important to consider that some of the indices of social functioning employed in the current study may have been inappropriate given the nature of the detention setting. For example, the anchors for low scores on the IRS correspond to youth not needing any treatment or special services and the anchors for high scores correspond to youth definitely needing treatment or special services. These anchors may be difficult for raters to apply in the detention setting as the detention setting may be conceptualized as a form of special service. As such, following this line of reasoning, if an adolescent is already receiving a special service, determining whether or not he needs special services may be a nonsensical task. Similarly, the social skills items on the SSIS measure prosocial

behaviors such as being polite. These types of behaviors may be less relevant in a social context wherein two primary competing social pressures are to adhere to facility rules and to fit in with peers who are also simultaneously trying to adhere to facility rules but also have a history of delinquent behavior. As such, the detention setting may represent a unique social context to which some of the items of the measures employed in this study do not apply. As such, it may be that it is neither the case that these measures are not sensitive to treatment effects nor that the treatment cannot change what it is intended to change. Rather, it may be that the constructs that these measures are intended to index are appropriate but the item content and wording may need to be adapted to better reflect the manifestations of social functioning that are relevant to the detention setting.

Limitations and Future Directions

Throughout the discussion, limitations of the current study have been highlighted. These included the small size of the sample, participant attrition, and that some of the measurements were not controlled. Data was also not obtained on the amount of time that elapsed between the time youth were admitted to the facility and the time they received ISG-CM. The amount of time between admittance and the treatment may influence youths' response to treatment. Conceivably, youth may be more compliant when admitted (e.g., due to not yet feeling comfortable with facility rules and staff) or close to discharge (e.g., due to aiming to preserve their good standing with their probation officer) than during the middle portion of their stay (e.g., due to feeling comfortable but not yet experiencing significant motivation to keep their discharge date). In addition to the recommendations articulated thus far, future evaluations of ISG-CM may benefit from

considering aggression and anger as among primary outcomes of interest as well as having long-term follow-up assessments of youths' functioning at the facility following ISG-CM. In addition, identifying characteristics that predict treatment response among juvenile offenders may be beneficial, especially given that this group of youth is highly heterogeneous. This would ultimately allow for youth to be prescribed treatments or treatment elements as a result of which they are most likely to benefit (i.e., a step towards personalized medicine). Ultimately, mechanisms of change in ISG-CM and necessary and sufficient components of ISG-CM should also be identified. Finally, given the long-term goal of the DFM process of dissemination, future trials of ISG-CM should include consultation, wherein researchers train facility staff to deliver ISG-CM independently and with adherence and fidelity.

Conclusion

The findings of this study demonstrate that ISG-CM was well accepted by this sample of detained juvenile offenders with externalizing disorders. The findings also demonstrate the feasibility of implementing ISG with juvenile offenders by incorporating adaptations for ED and the detention facility setting. Although preliminary, effectiveness data indicate ISG-CM was associated with pre- to post-treatment changes, relative to waitlist, in multiple domains including accuracy of self-appraisal of behavior, aggression, and daily behavior and participation at the facility and in facility activities. As such, these findings provide evidence of satisfaction with and of feasibility and effectiveness of ISG-CM with a population of youth who exhibit serious problems that adversely affect the

long-term futures of the youth, their families, and society; thus more development and evaluation work is needed.

REFERENCES

Adrian, M., Zeman, J., Erdley, C., Lisa, L., Homan, K., & Sim, L. (2009). Social contextual links to emotion regulation in an adolescent psychiatric inpatient population: Do gender and symptomatology matter? *Journal of Child Psychology and Psychiatry, 50,* 1428-1436.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

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

Axelrod, S. R., Perepletchikova, F., Holtzman, K., & Sinha, R. (2011). Emotion regulation and substance use frequency in women with substance dependence and borderline personality disorder receiving dialectical behavior therapy. *The American Journal of Drug and Alcohol Abuse*, *37*(1), 37-42.

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report: The Kentucky inventory of mindfulness skills. *Assessment, 11,* 191–206.

Ball, T. M., Stein, M. B., & Paulus, M. P. (2014). Toward the application of functional neuroimaging to individualized treatment for anxiety and depression. *Depression and Anxiety*, *31*(11), 920-933.

Barkley, R. A. (2010). Deficient emotional self-regulation: A core component of attention-deficit/hyperactivity disorder. *Journal of ADHD and Related Disorders, 1,* 5-37.

Beauchaine, T. P., Gatzke-Kopp, L., & Mead, H. K. (2007). Polyvagal theory and developmental psychopathology: Emotion dysregulation and conduct problems from preschool to adolescence. *Biological Psychology*, *74*(2), 174-184.

Becker, B. J. (1988). Synthesizing standardized mean change measures. *British Journal of Mathematical and Statistical Psychology, 41,* 257–278.

Bradley, B., DeFife, J. A., Guarnaccia, C., Phifer, J., Fani, N., Ressler, K. J., & Westen, D. (2011). Emotion dysregulation and negative affect: Association with psychiatric symptoms. *Journal of Clinical Psychiatry*, *72*(5), 685-691.

Brezina, T., & Piquero, A. R. (2007). Moral beliefs, isolation from peers, and abstention from delinquency. *Deviant Behavior, 28*, 433-465.

Brown, K.W., & Ryan, R.M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822–848.

Bunford, N. & Evans, S. W. (2013). *Parent-rated emotional symptoms differentiates youth with ISG mastery and non-mastery.* Unpublished manuscript, Department of Psychology, Ohio University, Athens, OH.

Bunford, N. & Evans, S. W. (2015). *The association between ODD and CD with self-report emotion dysregulation.* Unpublished manuscript, Department of Psychology, Ohio University, Athens, OH.

Bunford, N., Evans, S. W., Becker, S. A., & Langberg, J. M. (2015). Attention-deficit/hyperactivity disorder and social skills in youth: A moderated mediation

model of emotion dysregulation and depression. *Journal of Abnormal Child Psychology, 43,* 283-296.

Bunford, N., Evans, S. W., & Langberg, J. M. (2014). Emotion dysregulation is associated with social impairment among young adolescents with attention-deficit/hyperactivity disorder. *Journal of Attention Disorders.* DOI: 10.1177/1087054714527793

Bunford, N., Evans, S. W., Zoccola, P., & Rybak, T. (June, 2013). Poster presented at the International Society for Research in Child and Adolescent Psychopathology, Leuven, Belgium.

Bushman, B. J., Baumeister, R. F., & Phillips, C. M. (2001). Do people aggress to improve their mood? Catharsis beliefs, affect regulation opportunity, and aggressive responding. *Journal of Personality and Social Psychology*, *81*(1), 17-32.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63*(3), 452.

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*(3), 310-334.

Chen, X., & Adams, M. (2010). Are teen delinquency abstainers social introverts?: A test of Moffitt's theory. *Journal of Research in Crime and Delinquency, 47*, 439-468.

Coccaro, E. F., Harvey, P. D., Kupsaw-Lawrence, E., Herbert, J. L., & Bernstein, D. P. (1991). Development of neuropharmacologically based behavioral assessments of impulsive aggressive behavior. *Journal of Neuropsychiatry, 3*(2), S44–S51.

Dadds, M. R., Perry, Y., Hawes, D. J., Merz, S., Riddell, A. C., Haines, D. J., ... & Abeygunawardane, A. I. (2006). Attention to the eyes and fear-recognition deficits in child psychopathy. *The British Journal of Psychiatry*,*189*(3), 280-281.

Davidson, R. J., Putnam, K. M., & Larson, C. L. (2000). Dysfunction in the neural circuitry of emotion regulation--a possible prelude to violence. *Science*, *289*(5479), 591-594.

Dembo, R., Schmeidler, J., Pacheco, K., Cooper, S., & Williams, L. W. (1977). The relationships between youth's identified substance use, mental health or other problems at a juvenile assessment center and their referrals to needed services. *Journal of Child & Adolescent Substance Abuse, 6,* 23–54.

Desai, R. A., Goulet, J. L., Robbins, J., Chapman, J. F., Migdole, S. J., & Hoge, M. A. (2006). Mental health care in juvenile detention facilities: A review. *Journal of the American Academy of Psychiatry and the Law Online*, *34*(2), 204-214.

Downs, A., & Smith, T. (2004). Emotional understanding, cooperation, and social behavior in high-functioning children with autism. *Journal of Autism and Developmental Disorders*, *34*(6), 625-635.

Elliott, S. N., & Gresham, F. M. (2008). *SSIS-RS intervention guide*. Minneapolis, MN: NCS Pearson.

Evans, S. W., Brady, C. E., Harrison, J. R., Bunford, N., Kern, L., State, T. & Andrews, C. (2013). Measuring ADHD and ODD symptoms and impairment using high school teachers' ratings. *Journal of Clinical Child and Adolescent Psychology*, *42*(2), 197-207.

Evans, S. W., Owens, J. S., & Bunford, N. (2014b). Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. *Journal of Clinical Child & Adolescent Psychology*, *43*(4), 527-551.

Evans, S. W., Schultz, B. K., & DeMars, C. E. (2014a). High school based treatment for adolescents with ADHD: Results from a pilot study examining outcomes and dosage. *School Psychology Review*, *43*, 185–202.

Evans, S. W., Schultz, B. K., DeMars, C. E., Davis, H. (2011). Effectiveness of the Challenging Horizons after-school program for young adolescents with ADHD. *Behavior Therapy, 42*, 462–474.

Evans, S. W., Langberg, J. M., Schultz, B. K., Vaughn, A., Altaye, M., Marshall, S. A., & Zoromski, A. K. (2016). Evaluation of a school-based treatment program for young adolescents with ADHD. *Journal of Consulting and Clinical Psychology*, *84*(1), 15-30.

Fabiano, G. A., Pelham, W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M. et al. (2006). A practical measure of impairment: Psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child and Adolescent Psychology, 35*, 369-385.

Fiorentine, R., Nakashima, J., & Anglin, M. D. (1999). Client engagement in drug treatment. *Journal of Substance Abuse Treatment*, *17*(3), 199-206.

Friedman, S. R., Rapport, L. J., Lumley, M., Tzelepis, A., VanVoorhis, A., Stettner, L., & Kakaati, L. (2003). Aspects of social and emotional competence in adult attention-deficit/hyperactivity disorder. *Neuropsychology*, *17*(1), 50-58.

Fristad, M. A., Teare, M., Weller, E. B., Weller, R. A., & Salmon, P. (1998). Study III: Development and concurrent validity of the Children's Interview for Psychiatric Syndromes—parent version (P-ChIPS). *Journal of Child and Adolescent Psychopharmacology*, *8*(4), 221-226.

Furlong, M. J., & Smith, D. C. (1994). *Anger, hostility, and aggression: Assessment, prevention, and intervention strategies for youth*. Brandon, VT: Clinical Psychology Publishing Company, Inc.

Gibbons, R. D., Hedeker, D. R., & Davis, J. M. (1993). Estimation of effect size from a series of experiments involving paired comparisons. *Journal of Educational Statistics, 18,* 271–279.

Goldstein, T. R., Axelson, D. A., Birmaher, B. &. Brent, D. A. (2007). Dialectical behavior therapy for adolescents with bipolar disorder: A 1-year open trial. *American Academy of Child and Adolescent Psychiatry, 46*, 820-830.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41-54.

Greene, R. W., Biederman, J., Zerwas, S., Monuteaux, M. C., Goring, J. C., & Faraone, S. V. (2002). Psychiatric comorbidity, family dysfunction, and social impairment in referred youth with oppositional defiant disorder. *American Journal of Psychiatry, 159,* 1214-1224.

Gresham, F. M., Elliott, S. N., Vance, M. J., & Cook, C. R. (2011). Comparability of the Social Skills Rating System to the Social Skills Improvement System: Content and psychometric comparisons across elementary and secondary age levels. *School Psychology Quarterly, 26*, 27-44.

Guerra, N. G., & Slaby, R. G. (1990). Cognitive mediators of aggression in adolescent offenders: II. Intervention. *Developmental Psychology*, *26*(2), 269.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92,* 490–499.

Hinshaw, S. P. (2003). Impulsivity, emotion regulation, and developmental psychopathology: Specificity versus generality of linkages. *Annals of the New York Academy of Sciences, 1008*, 149-159.

Hirschi, T. (1969). *Causes of delinquency.* Berkeley, CA: University of California Press.

Holman, B., & Zeidenburg, J. (2013). Dangers of detention: The impact of incarcerating youth in detention and other secure facilities. *National Criminal Justice Reference Service.* Retrieved from https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=269394

Homack, S. & Reynolds, C. (2007). *Essentials of assessment with brief intelligence tests*. Hoboken, NJ: John Wiley & Sons, Inc.

Institute for Education Sciences (n.d.). 2012 National board for education sciences annual report briefing material for board members: The NCER goal structure. Retrieved from https://ies.ed.gov/director/board/briefing/ncer_structure.asp

Kats-Gold, I., Besser, A., Priel, B. (2007). The role of simple emotion recognition skills among school aged boys at risk of ADHD. *Journal of Abnormal Child Psychology, 35*, 363-78.

Kern, L., Evans, S. W., Lewis, T. J., Weist, M. D., & Wills, H. P. (2015). CARS comprehensive intervention for secondary students with emotional and behavioral problems conceptualization and development. *Journal of Emotional and Behavioral Disorders*, *23,* 195-205.

Kirchner, L. (2014, December 11). The cost of juvenile detention. *Pacific Standard.* Retrieved from http://www.psmag.com/business-economics/the-cost-of-juvenile-incarceration-96258

Li, Z., Snieder, H., Su, S., Ding, X., Thayer, J. F., Treiber, F. A., …, & Wang, X..(2009). A longitudinal study in youth of heart rate variability at rest and in response to stress. *International Journal of Psychophysiology, 73,* 212-217.

Linehan, M. M. (1993). *Skills Training Manual for Treating Borderline Personality Disorder*. New York, NY: Guilford.

Linehan, M. M., Bohus, M., & Lynch, T. R. (2007). Dialectical behavior therapy for pervasive emotion dysregulation: Theoretical and practical underpinnings. In

Gross, J. J. (Ed). *Handbook of Emotion Regulation*. New York, NY: Guilford Press.

Lochman, J. E., & Lampron, L. B. (1986). Situational social problem-solving skills and self-esteem of aggressive and nonaggressive boys. *Journal of Abnormal Child Psychology*, *14*(4), 605-617.

Lorch, E. P., Milich, R., Astrin, C. C., & Berthiaume, K. S. (2006). Cognitive engagement and story comprehension in typically developing children and children with ADHD from preschool through elementary school. *Developmental Psychology*, *42*(6), 1206-1219.

Lorenz, A. R., & Newman, J. P. (2002). Deficient response modulation and emotion processing in low-anxious Caucasian psychopathic offenders: results from a lexical decision task. *Emotion*, *2*(2), 91.

MacDermott, S. T., Gullone, E., Allen, J. S., King, N. J., & Tonge, B. (2010). The emotion regulation index for children and adolescents (ERICA): A psychometric investigation. *Journal of Psychopathology and Behavioral Assessment*, *32*(3), 301-314.

Marziller, J. S., & Spence, S. H. (1981). Social skill training with adolescent male offenders. *Behavior Research and Therapy*, *19*, 349-368.

McCord, J., Widom, C. S., & Crowell, N. A. (Eds.). (2001). *Juvenile crime, juvenile justice*. Washington, DC: National Academy Press.

Mennin, D. S., Heimberg, R. G., Turk, C. L., & Fresco, D. M. (2004). Preliminary evidence for an emotion dysregulation model of generalized anxiety disorder. *Behaviour Research and Therapy, 43,* 1281-1310.

Merikangas, K. R., Nakamura, E. F., & Kessler, R. C. (2009). Epidemiology of mental disorders in children and adolescents. *Dialogues in Clinical Neuroscience*, *11*(1), 7-20.

Milich, R. (1994). The response of children with ADHD to failure: If at first you don't succeed, do you try, try again? *School Psychology Review, 23*, 11–28.

Molina, B. S., Pelham, W. E., Blumenthal, J., & Galiszewski, E. (1998). Agreement among teachers' behavior ratings of adolescents with a childhood history of attention deficit hyperactivity disorder. *Journal of Clinical Child Psychology*, *27*(3), 330-339.

Nelson-Gray, R., Keane, S., Hurst, R., Mitchell, J., Warburton, J., Chok, J., & Cobb, A. (2006). A modified DBT skills training for oppositional defiant adolescents: Promising preliminary findings. *Behavior Research and Therapy, 44,* 1811-1820.

Nguyen, T. D., Attkisson, C. C., & Stegner, B. L. (1983). Assessment of patient satisfaction: development and refinement of a service evaluation questionnaire. *Evaluation and Program Planning*, *6*(3), 299-313.

Owens, J. G., & Slocum, L. A. (2015). Abstainers in adolescence and adulthood exploring the correlates of abstention using Moffitt's developmental taxonomy. *Crime & Delinquency, 61*, 690-718.

Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*(2), 210-218.

Pillow, D., Pelham, W. E., Hoza, B., Molina, B., & Stultz, C. (1998). Confirmatory factor analyses examining ADHD symptoms and other childhood disruptive behaviors. *Journal of Abnormal Child Psychology, 26*, 293-309.

Pulkkinen, L., & Pitkänen, T. (1993). Continuities in aggressive behavior from childhood to adulthood. *Aggressive Behavior*, *19*(4), 249-263.

Ross, R. R., & McKay, H. B. (1976). A study of institutional treatment programs. *International Journal of Offender Therapy and Comparative Criminology, 20,* 165-173.

Rottenberg, J., Kasch, K. L., Gross, J. J., & Gotlib, I. H. (2002). Sadness and amusement reactivity differentially predict concurrent and prospective functioning in major depressive disorder. *Emotion, 2*, 135–146.

Sadler, J. M., Evans, S. W., Schultz, B. K., & Zoromski, A. Z. (2011). Potential Mechanisms of Action in the Treatment of Social Impairment and Disorganization in Adolescents with ADHD. *School Mental Health, 3,* 156-168.

Schlichter, K. J., & Horan, J. J. (1981). Effects of stress inoculation on the anger and aggression management skills of institutionalized juvenile delinquents. *Cognitive Therapy and Research*, *5*(4), 359-365.

Schultz, B. K., Evans, S. W., Langberg, J. M., & Schoemann, A. M. (in preparation). Outcomes for adolescents who comply with long-term ADHD treatment.

Schultz, D., Izard, C. E., & Ackerman, B. P. (2000). Children's anger attribution bias: Relations to family environment and social adjustment. *Social Development*, *9*(3), 284-301.

Shelton, D., Kesten, K., Zhang, W., & Trestman, R. (2011). Impact of a dialectic behavior therapy—Corrections Modified (DBT-CM) upon behaviorally challenged incarcerated male adolescents. *Journal of Child and Adolescent Psychiatric Nursing*, *24*(2), 105-113.

Shivrattan, J. L. (1988). Social interactional training and incarcerated juvenile delinquents. *Canadian Journal of Criminology*, *30*, 145.

Slaby, R. G., & Guerra, N. G. (1988). Cognitive mediators of aggression in adolescent offenders: I. Assessment. *Developmental Psychology*, *24*(4), 580-588.

Sibley, M. H., Altszuler, A. R., Morrow, A. S., & Merrill, B. M. (2014). Mapping the academic problem behaviors of adolescents with ADHD. *School Psychology Quarterly, 29,* 422-437.

Surís, A. M., Lind, L. M., Kashner, M. T., Bernstein, I. H., Young, K., & Worchel, J. (2005). Aggression and impulsivity instruments: An examination in veterans. *Military Psychology*, *17*(4), 283-297.

Suveg, C., Morelen, D., Brewer, G. A., & Thomassin, K. (2010). The emotion dysregulation model of anxiety: A preliminary path analytic examination. *Journal of Anxiety Disorders, 24,* 924-930.

Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-Aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV–heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, *113*(1), 210-220.

Task Force of the European Society of Cardiology, & Task Force of the European Society of Cardiology. (1996). the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation*, *93*(5), 1043-1065.

Telch, C. F., Agras, W. S., & Linehan, M. M. (2001). Dialectical behavior therapy for binge eating disorder. *Journal of Consulting and Clinical Psychology, 69*, 1061-1065.

Teplin, L. A., Abram, K. M., McClelland, G. M., Dulcan, M. K., & Mericle, A. A. (2002). Psychiatric disorders in youth in juvenile detention. *Archives of General Psychiatry*, *59*(12), 1133-1143.

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, *36*(2), 747-756.

Thayer, J. F., & Sternberg, E. M. (2010). Neural aspects of immunomodulation: focus on the vagus nerve. *Brain, Behavior, and Immunity*, *24*(8), 1223-1228.

Timmons-Mitchell, J., Brown, C., Schulz, S. C., Webster, S. E., Underwood, L. A., Semple, W. E. (1997). Comparing the mental health needs of female and male incarcerated juvenile delinquents. *Behavioral Sciences & the Law, 15*, 195–202.

Trupin, E. W., Stewart, D. G., Beach, B., & Boesky, L. (2002). Effectiveness of a dialectical behaviour therapy program for incarcerated female juvenile offenders. *Child and Adolescent Mental Health*, *7*(3), 121-127.

Van Eck, K., Finney, S. J., & Evans, S. W. (2010). Parent Report of ADHD Symptoms of Early Adolescents: A Confirmatory Factor Analysis of the Disruptive Behavior Disorders Scale. *Educational & Psychological Measurement, 70*(6), 1042-1059.

Vasilev, C. A., Crowell, S. E., Beauchaine, T. P., Mead, H. K., & Gatzke-Kopp, L. M. (2009). Correspondence between physiological and self-report measures of emotion dysregulation: A longitudinal investigation of youth with and without psychopathology. *Journal of Child Psychology and Psychiatry, 50*, 1357–1364.

Walker, J. L., Lahey, B. B., Russo, M. F., Frick, P. J., Christ, M. A. G., McBurnett, K., ... & Green, S. M. (1991). Anxiety, inhibition, and conduct disorder in children: I. Relations to social impairment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *30*(2), 187-191.

Watson, P. (2016, January 14). Rules of thumb on magnitudes of effect sizes. Retrieved from: http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize

Webster-Stratton, C. (1996). Early-onset conduct problems: Does gender make a difference?. *Journal of Consulting and Clinical Psychology*, *64*(3), 540-551.

Wechsler D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI).* San Antonio, TX: The Psychological Corporation.

Weinberg, A., & Klonsky, E. D. (2009). Measurement of emotion dysregulation in adolescents. *Psychological Assessment, 21*, 616-621.

Weisz, J. R., Jensen, A. L., & McLeod, B. D. (2005). Development and dissemination of child and adolescent psychotherapies: Milestones, methods, and a new deployment-focused model. In E. D. Hibbs & P. S. Jensen (Eds.). *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice* (2nd ed., pp. 9–39). Washington, DC: American Psychological Association.

Weller, E. B., Weller, R. A., Fristad, M. A., Rooney, M. T., & Schecter, J. (2000). Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 76–84.

Zuddas, A., Marzocchi, G., Oosterlaan, J., Cavolina, P., Ancilletta, B., & Sergeant, J. (2006). Factor structure and cultural factors of disruptive behavior disorders symptoms in Italian children. *European Psychiatry, 21*, 410-418.

Table 1 *Modifications to the Original ISG Protocol and Corresponding Rationale*

Modification	Rationale
The Psychoeducation phase of ISG was expanded to include education about: (i) emotions (e.g., basic emotions, nonverbal and verbal expression of emotions, the distinction between primary and secondary emotions, and the utility of emotions in communication and motivation),	Youth with externalizing disorders have deficiencies in their ability to identify and understand their own and others' emotions (Dadds et al., 2006; Downs & Smith, 2004; Kats-Gold, Besser, & Priel, 2007). In turn, deficient emotion recognition and understanding is associated with problematic development of social skills (Friedman et al., 2003) and exacerbates extant social problems (Kats-Gold et al., 2007). It was hypothesized that education on emotions may increase youths' ability to label their emotions and therefore better identify and understand them.
(ii) emotion regulation, with particular focus on aspects of ED that are associated with social impairment in adolescents with externalizing symptoms (Bunford, et al., 2014; 2015), and the implications of emotion regulation for developing and maintaining relationships with others; and (iii) mindfulness meditation	Psychoeducation is a common part of psychosocial treatments (e.g., Evans, Schultz, DeMars, & Davis, 2011). Psychoeducation is important for participants to learn the vocabulary of the treatment and also necessary for communication, skill building, and learning effective steps of problem-solving. Thus, it was important for youth to learn vocabulary related to ISG-CM, including the key concept of emotion regulation and the regular practice of mindfulness meditation.
The problem solving steps of the Psychoeducation phase were modified to include a direct focus on emotions. e.g., "What is my problem" has been replaced by "What prompted my emotion and what is my emotion"; "List some ideas I could do" has been replaced by "List some ideas I could do if my emotion is negative" and "List some ideas I could do if my emotion is positive"	Youth with ADHD have difficulties attending to their emotional states and accepting and acknowledging those when they occur (Bunford, et al., 2014). Aggressive youth exhibit an anger or hostile attribution bias in their social problem solving (see above). A relevant technique in DBT targets difficulties with accepting and acknowledging emotions as well as increases non-emotional problem solving. Participants are trained to identify and label their emotions by describing (1) the event that elicited the emotional reaction, (2) their interpretations of such event, (3) the phenomenological experience of the emotion (including the accompanying physical sensations), (4) the behaviors used to express the emotion, and (5) the effects or aftereffects of the emotion on other areas of functioning (Linehan, 1993). Aspects of this technique were incorporated to modify the original ISG problem solving steps and hypothesized to improve acceptance of emotions and non-emotional (including non-angry or hostile problem solving).

Table 1 continued

Although in traditional ISG there are no restrictions with regard to Ideal Self Goals (other than those needing to be appropriate for the purpose and setting), in ISG-CM (iv) at least 50% of a youth's Ideal Self Goals and their operational definitions are to relate to emotion regulation[7].	The purpose of these changes was to increase direct focus on emotion regulation. In addition, prior findings indicated, at least with youth with ADHD, that emotion regulation is a mechanism through which ADHD impacts social functioning (Bunford et al., 2015). Thus, it was expected that it is necessary to improve ED to achieve gains in many areas of social functioning.
(v) Feedback sessions during practice activities have been modified to include self-ratings of emotional arousal on an Emotions Thermometer (see Appendix)	Youth with externalizing disorders have deficiencies in their ability to identify and understand emotions (Dadds et al., 2006; Downs & Smith, 2004; Kats-Gold, et al., 2007). By having youth repeatedly identify their emotion and the intensity thereof, the goal was to train youth to routinely attend to their emotional state (and therefore prevent secondary emotions, the most common of which is anger [Linehan, 1993]. Because anger is a precursor to aggression, a characteristic that confers risk for criminal offending, we expected that minimizing the occurrences of anger as a secondary emotion would result in benefits to social functioning).
and a discussion of the link between this arousal and difficulty/ease in behaving congruently with Ideal Self Goals	Youth with ADHD (Bunford et al., 2014) and with ODD and CD (Bunford & Evans, 2016; unpublished manuscript) exhibit poor behavioral control when experiencing strong emotions, in particular with adapting outward behavioral manifestations of emotional states in a way that those are congruent with (social) situational demands. As such, training youth to attend to their emotional arousal and link such arousal to the relative degree to which their emotional states are congruent with social demands was hypothesized to improve their ability, over time, to manage the outward expressions of their emotions.
In traditional ISG the activities from which youth and group leaders choose are not arranged in a predetermined order. In ISG-CM, the (vi) activities are arranged on a schedule that is to present	These changes were made given prior encouraging findings on in vivo provocations (Schlicter & Horan, 1981) as well as the intention to systematically expose youth to increasingly more difficult situations so as to ensure continued opportunity to learn and practice emotion regulation skills.

[7] e.g., "calm": "Even when others do, I do not show or feel nervousness/anger/other emotion", "I often feel relaxed", "It is easy for me to relax"; "patient": "I continue to talk in a friendly and normal pace even when others are excited", "When forced to wait for things, I don't complain", "I maintain a pleasant demeanor when am very happy"

Table 1 continued

increasingly more opportunities for frustration.	
Finally, (vii) youth and group leaders practice mindfulness meditation at the end of every group session and youth are encouraged to practice it outside of group.	Emotional reactivity is associated with ADHD (e.g., Barkley, 2010) and with conduct (Lorenz & Newman, 2002; Webster-Stratton, 1996) and externalizing problems. Mindfulness meditation decreases emotional reactivity (Baer, Smith, Allen, 2004; Brown & Ryan, 2003; Mennin, Heimberg, Turk, & Fresco, 2005; Rottenberg, Kasch, Gross, & Gotlib, 2002; Suveg, Morelen, Brewer, & Thomassin, 2010).

Table 2 *Symptoms and Diagnoses for Disorders Other than ADHD, ODD, and CD, Assessed via Self-Report at Baseline*

Youth	# of SAD symptoms	SAD	# of GAD symptoms	GAD	# of PTSD symptoms	PTSD	# of depression symptoms	MDD	# of SP symptoms	SP	SUBAB
1	0	No	0	No	0	No	0	No	0	No	No
2	0	No	3	Yes	11	Yes	1	No	3[a]	Situational type	No
3	0	No	0	No	7	No	4	No	0	No	Yes
4	0	No	0	No	14	Yes	9	Yes	3	No	No
5	5	Yes	2	No	0	No	6	Yes	0	No	No
6	0	No	3	Yes	0	No	12	Yes	4[b]	No	Yes
7	0	No	0	No	0	No	11	Yes	0	No	No
8	0	No	0	No	1	No	0	No	0	No	Yes
9	0	No	3	Yes	11	Yes	1	No	5[c]	Animal and Situational type	No
10	0	No	0	No	0	No	0	No	0	No	Yes
11	0	No	0	No	1	No	0	No	3[d]	No	Yes
12	0	No	0	No	0	No	0	No	3[e]	No	Yes
Total	-	1	-	3	-	3	-	4	-	2	6

Note. SAD = Social Anxiety Disorder, GAD = Generalized Anxiety Disorder, PTSD = Post-traumatic Stress Disorder, MDD = Major Depressive Disorder, SP = Specific Phobia, [a] = airplanes, heights, [b] = snakes, spiders, [c] = crossing bridges, elevators, fire, heights, snakes, spiders, thunderstorms, [d] = bugs, crowds, lightning, mice, needles, parrots, rats, snakes, spiders, thunderstorms, [e] = spiders. None of the youth reported any symptoms of Mania, Hypomania, Dysthymic Disorder, Psychosis/Schizophrenia, Obsessive-Compulsive Disorder, or Acute Stress Disorder.

Table 3 ADHD Symptoms, Related Impairment, and Diagnoses for Self- and Parent-Report at Baseline

Youth	# of ADHD-IA symptoms	# of ADHD-H/I symptoms	# of ADHD symptoms	Impairment			ADHD	# of parent-reported ADHD-IA symptoms on the DBD (diagnosis)	# of parent-reported ADHD-H/I symptoms on the DBD (diagnosis)	Subtype based on parent DBD
				Home	School	Peer				
1	5	2	7	No	No	Yes	No	8 (Yes)	7 (Yes)	Comb
2	4	1	5	No	No	No	No	7 (Yes)	9 (Yes)	Comb
3	5	2	7	No	Yes	No	No	0 (No)	0 (No)	N/A
4	0	2	2	No	No	No	No	Missing	Missing	Missing
5	7	1	8	Yes	Yes	No	Yes, IA	6 (Yes)	1 (No)	IA
6	7	4	11	Yes	Yes	No	Yes, IA	6 (Yes)	7 (Yes)	Comb
7	6	3	9	Yes	No	No	Yes, IA	0 (No)	0 (No)	N/A
8	5	6	11	No	Yes	No	Yes, H/I	0 (No)	0 (No)	N/A
9	5	6	11	Yes	Yes	Yes	Yes, H/I	5 (No)	8 (Yes)	H/I
10	7	7	14	No	Yes	No	Yes, Comb	1 (No)	0 (No)	N/A
11	6	6	12	Yes	Yes	No	Yes, Comb	7 (Yes)	3 (No)	IA
12	2	0	2	No	No	No	No	1 (No)	1 (No)	N/A
Total	-	-	-	-	-	-	7	5	4	6

Note. ADHD = Attention-Deficit/Hyperactivity Disorder, IA = Inattention, H/I = Hyperactivity/Impulsivity, Comb = Combined; DBD = Disruptive Behavior Disorders Rating Scale.

Table 4 ODD Symptoms, Related Impairment, and Diagnoses for Self- and Parent-Report at Baseline

Youth	# of ODD symptoms	ODD	Home impairment	School impairment	Peer impairment	# of parent-reported ODD symptoms on the DBD (diagnosis)
1	4	Yes	Yes	No	No	7 (Yes)
2	5	Yes	Yes	Yes	Yes	8 (Yes)
3	1	No	No	No	No	0 (No)
4	0	No	No	No	No	Missing
5	3	No	No	No	No	6 (Yes)
6	5	Yes	Yes	Yes	Yes	6 (Yes)
7	1	No	No	No	No	0 (No)
8	4	Yes	Yes	Yes	No	2 (No)
9	4	Yes	No	Yes	Yes	8 (Yes)
10	5	Yes	No	Yes	No	0 (No)
11	6	Yes	Yes	Yes	Yes	3 (No)
12	2	No	No	No	No	2 (No)
Total	-	7	-	-	-	5

Note. ODD = Oppositional-defiant disorder, DBD = Disruptive Behavior Disorders Rating Scale.

Table 5 CD Symptoms, Related Impairment, and Diagnoses for Self- and Parent-Report at Baseline

Youth	# of CD symptoms	CD	Home impairment	School impairment	Peer impairment	Type	Severity	# of parent-reported CD symptoms on the DBD (diagnosis)
1	7	Yes	Yes	No	No	AO	Moderate	10 (Yes)
2	5	Yes	Yes	Yes	No	CO	Mild	9 (Yes)
3	8	Yes	No	No	No	CO	Moderate	1
4	4	Yes	No	No	No	AO	Mild	Missing
5	3	Yes	Yes	Yes	Yes	AO	Mild	3 (Yes)
6	6	Yes	Yes	Yes	No	AO	Moderate	1
7	5	Yes	Yes	Yes	No	AO	Moderate	0
8	9	Yes	Yes	Yes	No	AO	Moderate	0
9	9	Yes	Yes	Yes	Yes	AO	Severe	9 (Yes)
10	6	Yes	No	No	Yes	AO	Moderate	2
11	11	Yes	Yes	Yes	Yes	AO	Severe	9 (Yes)
12	9	Yes	Yes	Yes	No	AO	Moderate	3 (Yes)
Total	-	12	-	-	-	-	-	6

Note. CD = Conduct disorder, AO = Adolescent onset; CO = Childhood onset; DBD = Disruptive Behavior Disorders Rating Scale.

Table 6 Assessment Schedule for Measures Administered Across Groups and Time

	Group		
Measure	Grp1	Grp2-W	Grp2-T
ChIPS	T1	T1	
WASI-II	T1	T1	
DBD–Parent	T1	T1	
HRV	T1, T2		T2, T3
Ideal Self Goal ratings	During treatment		During treatment
DERS	T1, T2, T3	T1, T2, T3	
ERICA	T1, T2, T3	T1, T2, T3	
SSIS-RS	T1, T2, T3	T1, T2, T3	
BPAQ	T1, T2, T3	T1, T2, T3	
MOAS	T1, T2, T3	T1, T2, T3	
Facility behavior points	During treatment	During treatment	During treatment
SMR-Staff	During treatment		During treatment
SMR-Teacher	T1, T2, T3	T1, T2, T3	
DBD–Teacher	T1, T2, T3	T1, T2, T3	
IRS	T1, T2, T3	T1, T2, T3	
Satisfaction Questionnaire	T2		T3
PTEQ	T2		T3

Note. Grp1 = First treatment group, Grp2 = Second treatment group while waitlisted, Grp2-T = second treatment group while in treatment.
ChIPS = Children's Interview for Psychiatric Syndromes, WASI-II = Wechsler Abbreviated Scale of Intelligence, DBD-Parent = Disruptive Behavior Disorders Rating Scale-Parent report, HRV = heart rate variability, DERS = Difficulties in Emotion Regulation Scale, ERICA = Emotion Regulation Index for Children and Adolescents, SSIS-RS = Social Skills Improvement System-RS-Self report, BPAQ = Buss-Perry Aggression Questionnaire, MOAS = Modified Overt Aggression Scale, SMR-Staff = Social and Mood Ratings – Staff report, SMR-Teacher = Social and Mood Ratings – Teacher report, DBD-Teacher = Disruptive Behavior Disorders Rating Scale – Teacher report, IRS = Impairment Rating Scale, PTEQ = Preferred Treatment Elements Questionnaire.

Table 7 Descriptive Statistics for Satisfaction and Preferred Treatment Elements Self-Report Ratings

Satisfaction Questionnaire		
Total score	*n*	%
21.00	1	8.3
25.00	1	8.3
28.00	1	8.3
29.00	1	8.3
33.00	1	8.3
34.00	3	25.0
35.00	3	25.0
36.00	1	8.3
Missing	0	0

Preferred Treatment Elements Questionnaire						
	Liked					
	Psychoeducation		Group Sessions		Mindfulness Meditation	
Ranking	*n*	%	*n*	%	*n*	%
First	1	8.3	8	66.7	1	8.3
Second	3	25.0	1	8.3	6	50.0
Third	5	41.7	1	8.3	3	25.0
Missing	3	25.0	2	16.7	2	16.7
	Beneficial/Helpful					
First	6	50.0	5	55.6	1	8.3
Second	3	25.0	2	22.2	5	41.7
Third	9	75.0	1	11.1	3	25.0
Missing	3	25.0	4	33.3	3	25.0

Note. The total score on the Satisfaction Questionnaire is calculated by summing ratings (ranging from 1 to 4) across 9 items. The *n* and % correspond to the number and portion of youth who rated the treatment at the level of each of the obtained values.

Data on the Preferred Treatment Elements Questionnaire is presented for each of three treatment components, with corresponding rankings of first, second and third most liked and most beneficial/helpful rankings. The *n* and % correspond to the number and portion of youth who rated the treatment component at the level of each of the rankings.

.

Table 8 Group x Time Interactions

	Group x Time		
	F	*p*	η_p^2
Variable			
Emotion Regulation			
Teacher (2) SMR Emotional Regulation	2.461	.145	.183
Teacher (3) SMR Emotional Regulation	2.051	.180	.157
Social Functioning			
Staff SMR Social	1.865	.205	.172
Teacher SMR Social	3.062	.108	.218
Aggression			
BPAQ Verbal Aggression	1.425	.263	.137
MOAS Verbal Aggression	2.899	.119	.225
MOAS Property Aggression	2.151	.173	.177
ADHD Symptoms			
DBD-T Inattention	7.355	**.030**	.512
DBD-T H/I	.667	.460	.143
Academic Functioning			
IRS Academic Functioning	1.436	.261	.138
Externalizing/Internalizing Problems			
SSIS Internalizing	1.687	.223	.144

Note. SMR = Social and Mood Ratings; BPAQ = Buss-Perry Aggression Questionnaire; MOAS = Modified Overt Aggression Scale; ADHD = Attention-Deficit/Hyperactivity Disorder; H/I = Hyperactivity/Impulsivity; DBD-T = Disruptive Behavior Disorders Rating Scale-Teacher Version; IRS = Impairment Rating Scale; SSIS = Social Skills Improvement System Rating Scale

Table 9 Descriptive Statistics and Effect Size Estimates Comparing the Treatment and Waitlist Groups Pre- and Post-Treatment

	Pre-treatment *M (SD)*		Cohen's *D*	Post-treatment *M (SD)*		Cohen's *D*	|Repeated measures ES|
	T	W		T	W		
Emotion Regulation							
Teacher (2) SMR ER	12.29 (1.70)	11.67 (1.51)	0.38	12.43 (0.98)	11.33 (1.51)	0.86	0.31
Teacher (3) SMR ER	14.71 (5.50)	13.17 (5.85)	0.27	16.57 (2.94)	13.83 (3.06)	0.91	0.22
Social Functioning							
Staff SMR Social	7.46 (1.76)	7.25 (1.63)	0.12	8.21 (0.78)	6.65 (1.95)	1.05	0.80
Teacher SMR Social	7.14 (1.95)	7.17 (1.83)	-0.01	7.71 (1.89)	6.17 (2.04)	0.79	0.84
Aggression							
BPAQ Verbal	19.00 (4.60)	22.33 (9.11)	-0.46	19.43 (3.74)	25.00 (3.94)	-1.45	1.43
MOAS Verbal	1.00 (1.00)	0.67 (0.52)	0.42	0.43 (0.53)	1.80 (2.39)	-0.79	0.23
MOAS Property	0.67 (1.08)	0.10 (0.10)	0.75	0.39 (0.86)	1.30 (2.78)	-0.44	0.44
DBD Symptoms							
DBD-T Inattention	16.20 (11.30)	16.83 (6.68)	-0.07	12.17 (7.76)	16.00 (5.72)	-0.56	0.65
DBD-T H/I	16.33 (12.70)	14.75 (5.32)	0.16	11.33 (3.67)	15.00 (3.74)	-0.99	0.20

Table 9 continued

Academic Functioning							
IRS Academic Progress	3.86 (.73)	4.00 (.79)	-0.18	3.28 (.73)	4.5 (.79)	-1.60	2.77
Externalizing/Internalizing Problems							
SSIS Internalizing	9.33 (7.31)	6.17 (3.97)	0.54	6.14 (4.41)	7.00 (4.98)	-0.18	2.27

Note. T = Treatment Group; W = Waitlist-Control Group; ES = effect size; SMR = Social and Mood Ratings; ER = Emotion Regulation; BPAQ = Buss-Perry Aggression Questionnaire; MOAS = Modified Overt Aggression Scale; ADHD = Attention-Deficit/Hyperactivity Disorder; DBD-T = Disruptive Behavior Disorders Rating Scale-Teacher Version; H/I = Hyperactivity/Impulsivity; IRS = Impairment Rating Scale; SSIS = Social Skills Improvement System Rating Scale.

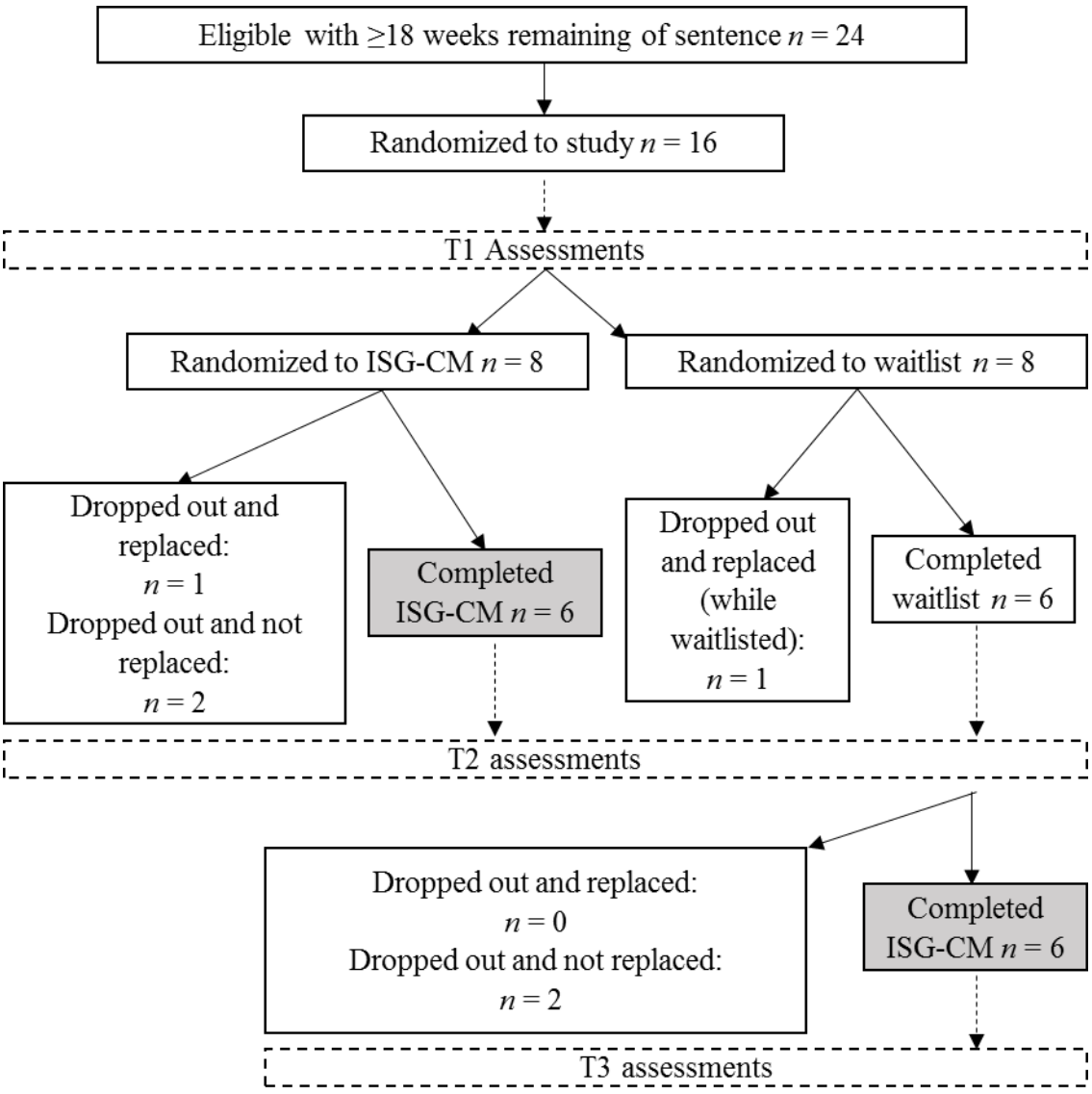
Pre-treatment Cohen's *D*s were calculated, in line with Hedges (1981; 1982) using the following formula: $(M_{\text{pre, treatment group}} - M_{\text{pre, waitlist group}})/(\text{SQRT}((SD_{\text{pre, treatment group}} * SD_{\text{pre, treatment group}} + SD_{\text{pre, waitlist group}} * SD_{\text{pre, waitlist group}})/2))$; post-treatment Cohen's *D*s were calculated using: $(M_{\text{post, treatment group}} - M_{\text{post, waitlist group}})/(\text{SQRT}((SD_{\text{post, treatment group}} * SD_{\text{post, treatment group}} + SD_{\text{post, waitlist group}} * SD_{\text{post, waitlist group}})/2))$; and repeated measures effect sizes were calculated, in line with Becker (1988) using: $((M_{\text{post, treatment group}} - M_{\text{pre, treatment group}})/SD_{\text{pre, treatment group}})) - ((M_{\text{post, waitlist group}} - M_{\text{pre, waitlist group}})/SD_{\text{pre, waitlist group}}))$.

Table 10 Descriptive Statistics and Effect Size Estimates Comparing T1, T2 and T3

	T1 *M (SD)*	T2 *M (SD)*	Repeated measures ES	T2 *M (SD)*	T3 *M (SD)*	Repeated measures ES
Emotion Regulation						
Impulse	20.60 (2.41)	19.40 (3.36)	-0.25	19.40 (3.36)	19.20 (0.84)	-0.06
Anger BPAQ	31.40 (8.17)	31.20 (6.30)	-0.02	31.20 (6.30)	19.00 (6.36)	-1.16
Social Functioning						
Overall impairment IRS	3.20 (2.17)	3.75 (0.96)	0.28	3.75 (0.96)	3.60 (0.89)	-0.45
Teacher-student relations impairment IRS	3.00 (2.65)	4.40 (2.51)	0.92	4.40 (2.51)	3.60 (2.30)	-0.25
Aggression						
BPAQ Verbal	22.60 (10.16)	25.00 (3.94)	0.26	25.00 (3.94)	10.00 (6.00)	-1.64
MOAS Verbal	0.60 (0.55)	1.80 (2.39)	0.56	1.80 (2.39)	1.60 (1.34)	-0.11
MOAS Property	0.10 (0.10)	1.30 (2.78)	0.45	1.30 (2.78)	1.70 (2.71)	0.11
Academic Functioning						
Classroom impairment IRS	2.80 (2.49)	3.50 (1.91)	0.10	3.50 (1.91)	2.60 (2.51)	-0.71

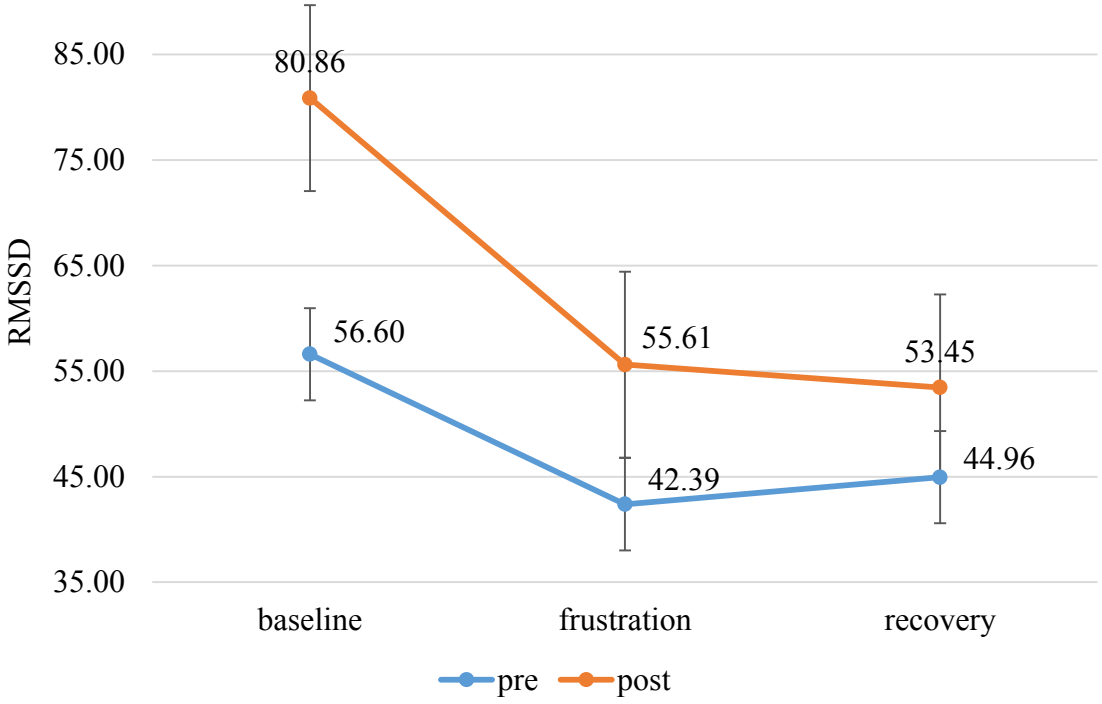
Note. ES = effect size; BPAQ = Buss-Perry Aggression Questionnaire; MOAS = Modified Overt Aggression Scale; IRS = Impairment Rating Scale.

Repeated measures effect sizes were calculated, in line with Gibbons, Hedeker, and Davis (1993) using the following formula: $(M_{T2} - M_{T1})/SD_{T2, T1}$) and $(M_{T3} - M_{T2})/SD_{T3, T2}$). Here, the numerator is the sample mean change, or the mean difference between T1 and T2 (and T2 and T3) scores, and the denominator is the sample standard deviation of change scores.



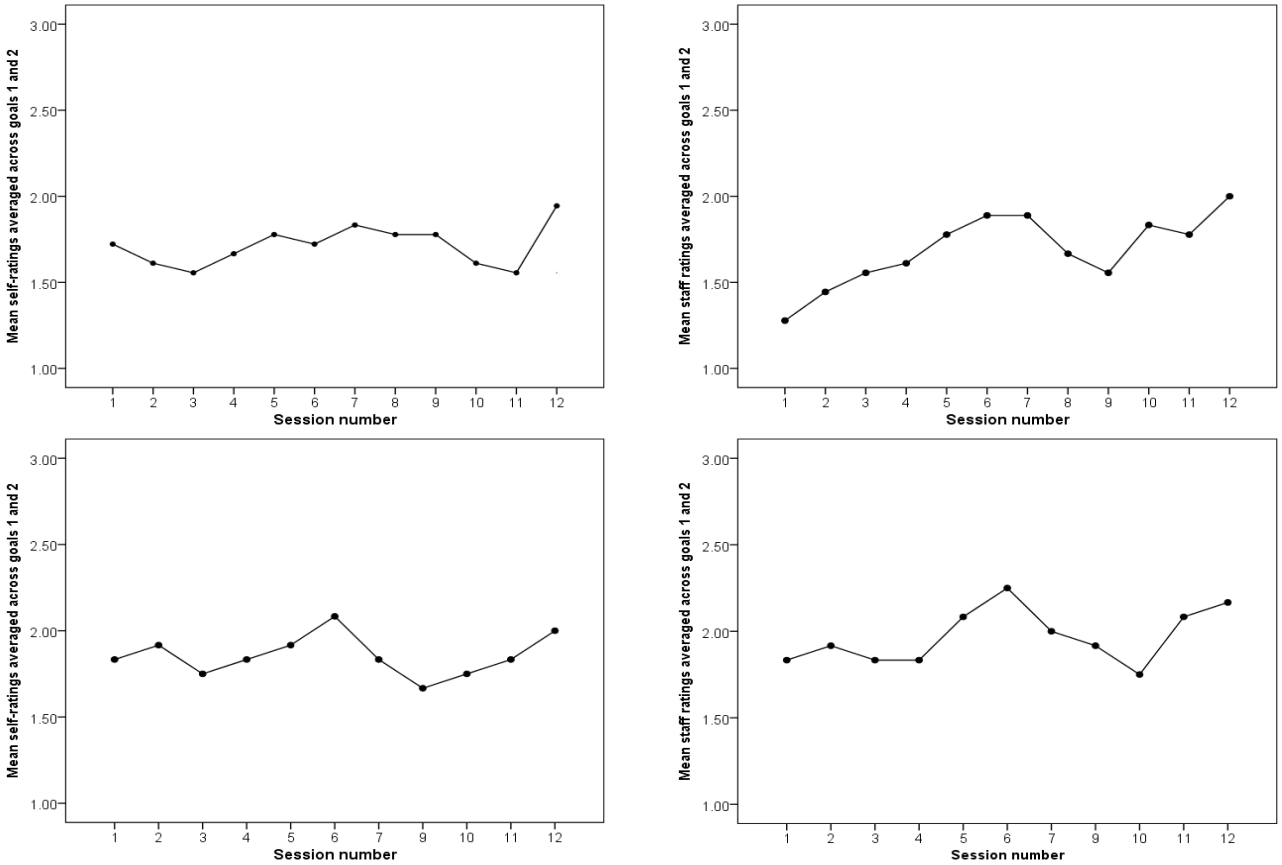
Note. ISG-CM = Interpersonal Skills Group - Corrections Modified.

Figure 1 Participant and assessment and treatment flow chart.



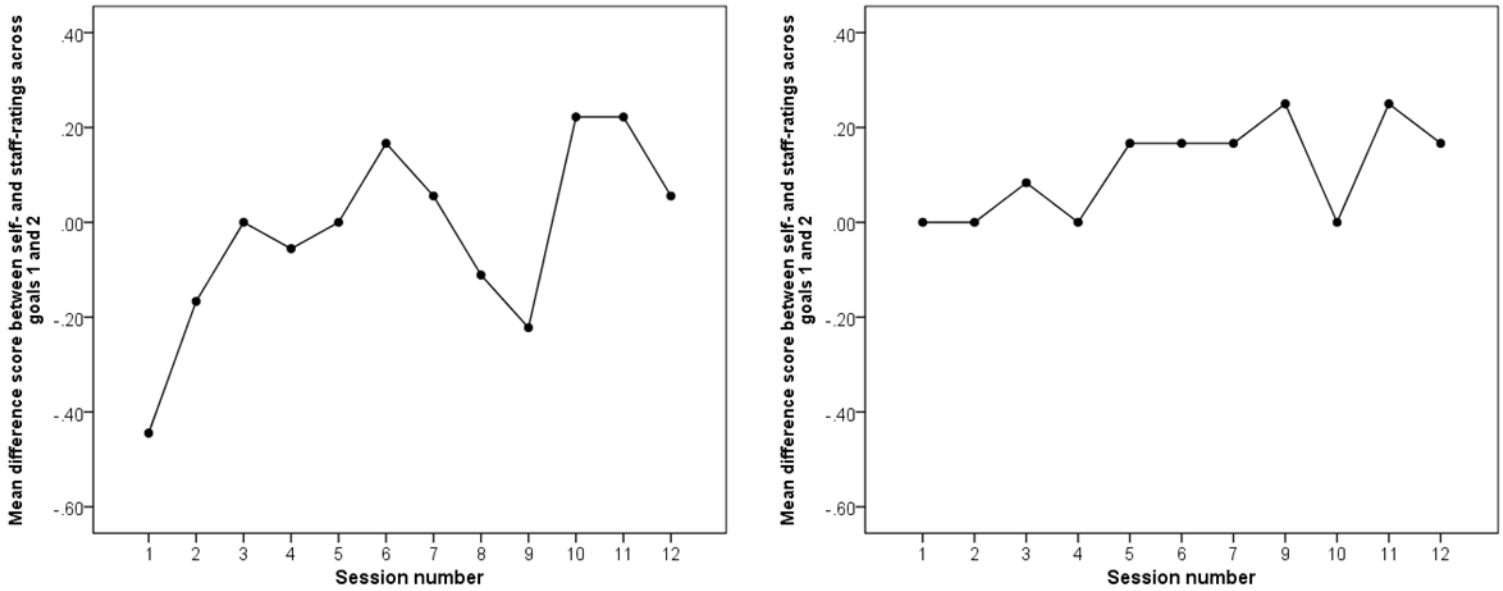
Note. Higher scores reflect better emotion regulation.

Figure 2 Pre- and post-treatment HRV during resting baseline, frustration task, and recovery.



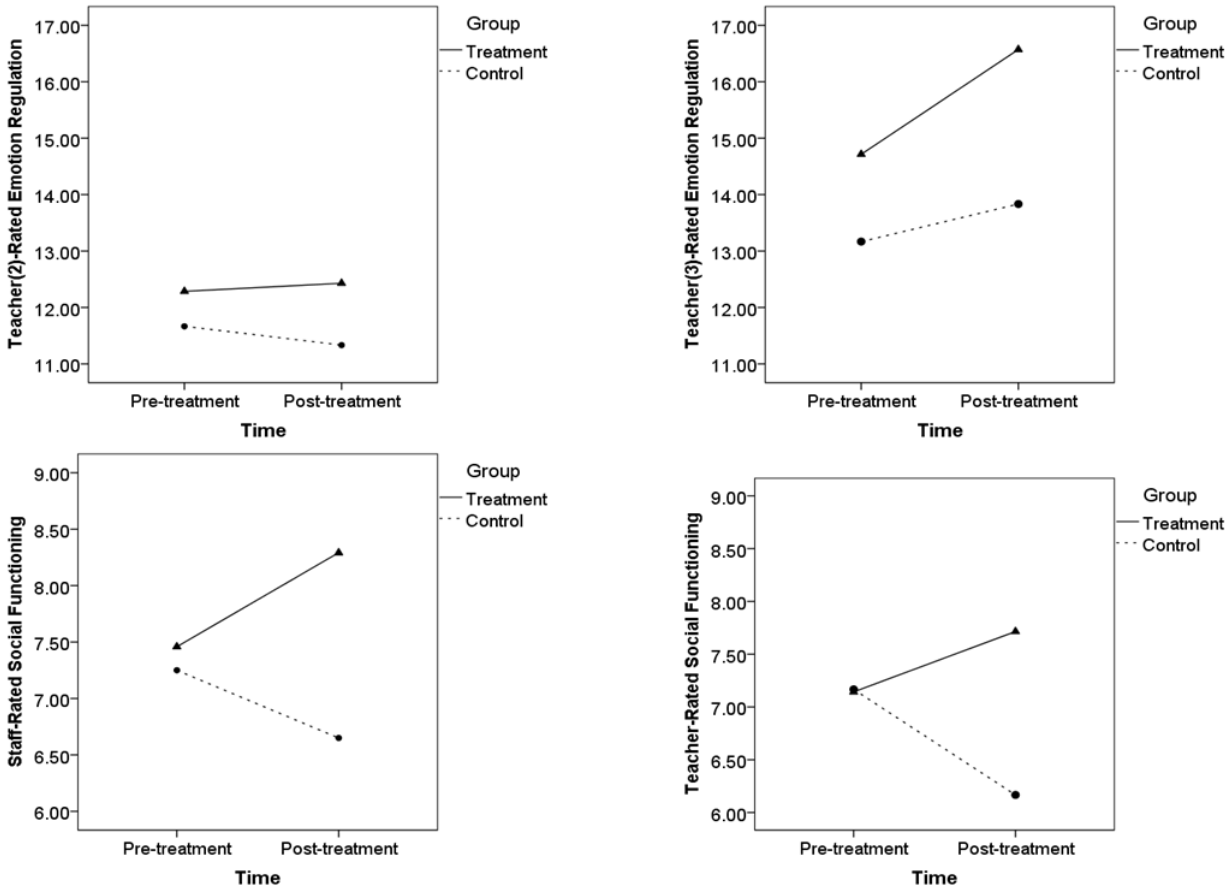
Note. Self-ratings are depicted on the left and clinician ratings are depicted on the right of Ideal Self Goal mastery for the first (depicted on top) and for the second feedback per session (depicted on bottom). No averages fell below 1. Higher scores indicate increased behavioral consistency with social goals.

Figure 3 Self- and clinician ratings of Ideal Self Goal mastery for the first feedback per session and for the second feedback per session.



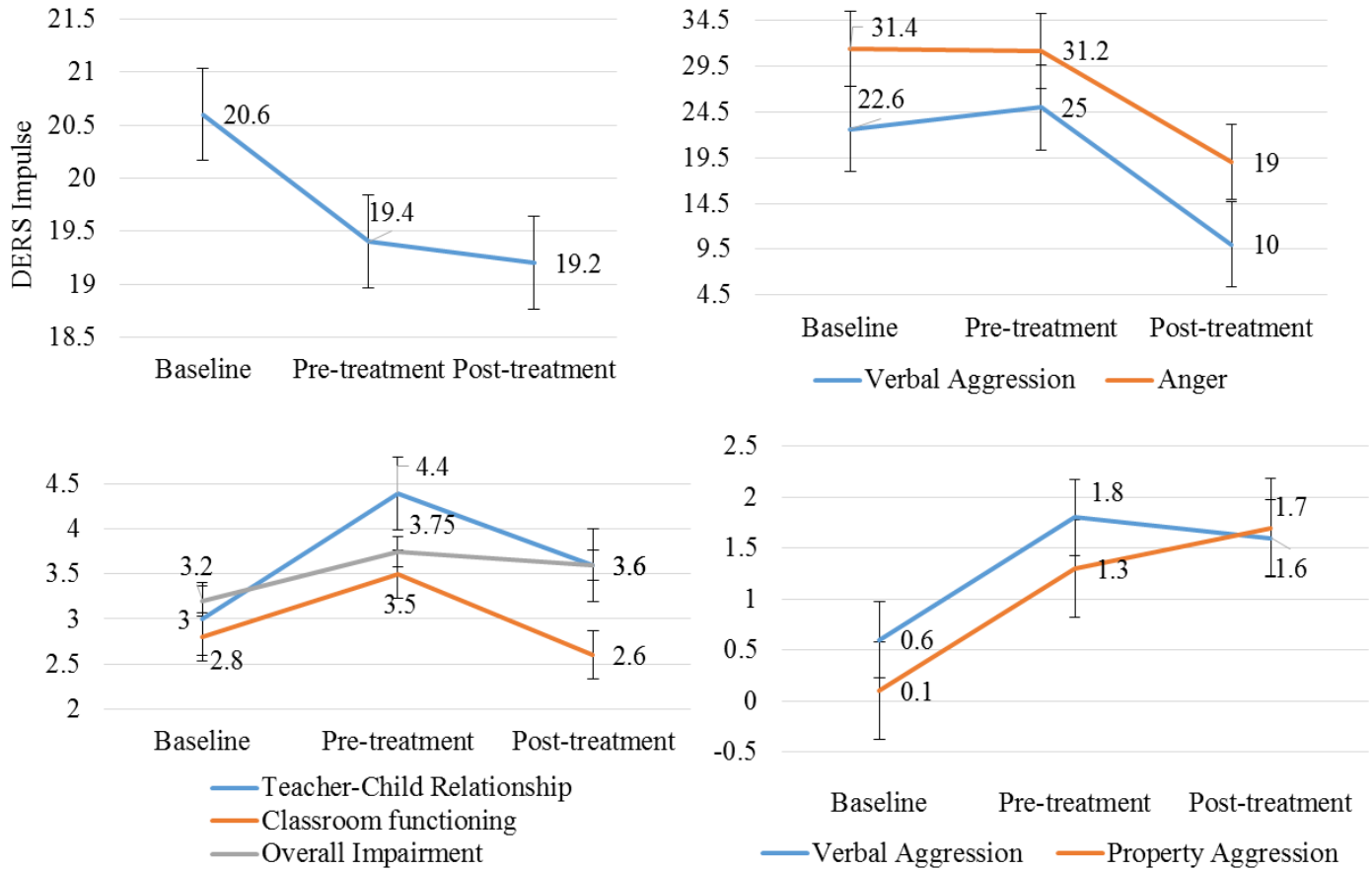
Note. Self-ratings are subtracted from clinician ratings for the first feedback per session (left) and for the second feedback per session (right). Higher scores indicate lower self-ratings compared to clinician ratings.

Figure 4 Difference scores between self- and clinician ratings of Ideal Self Goal mastery.



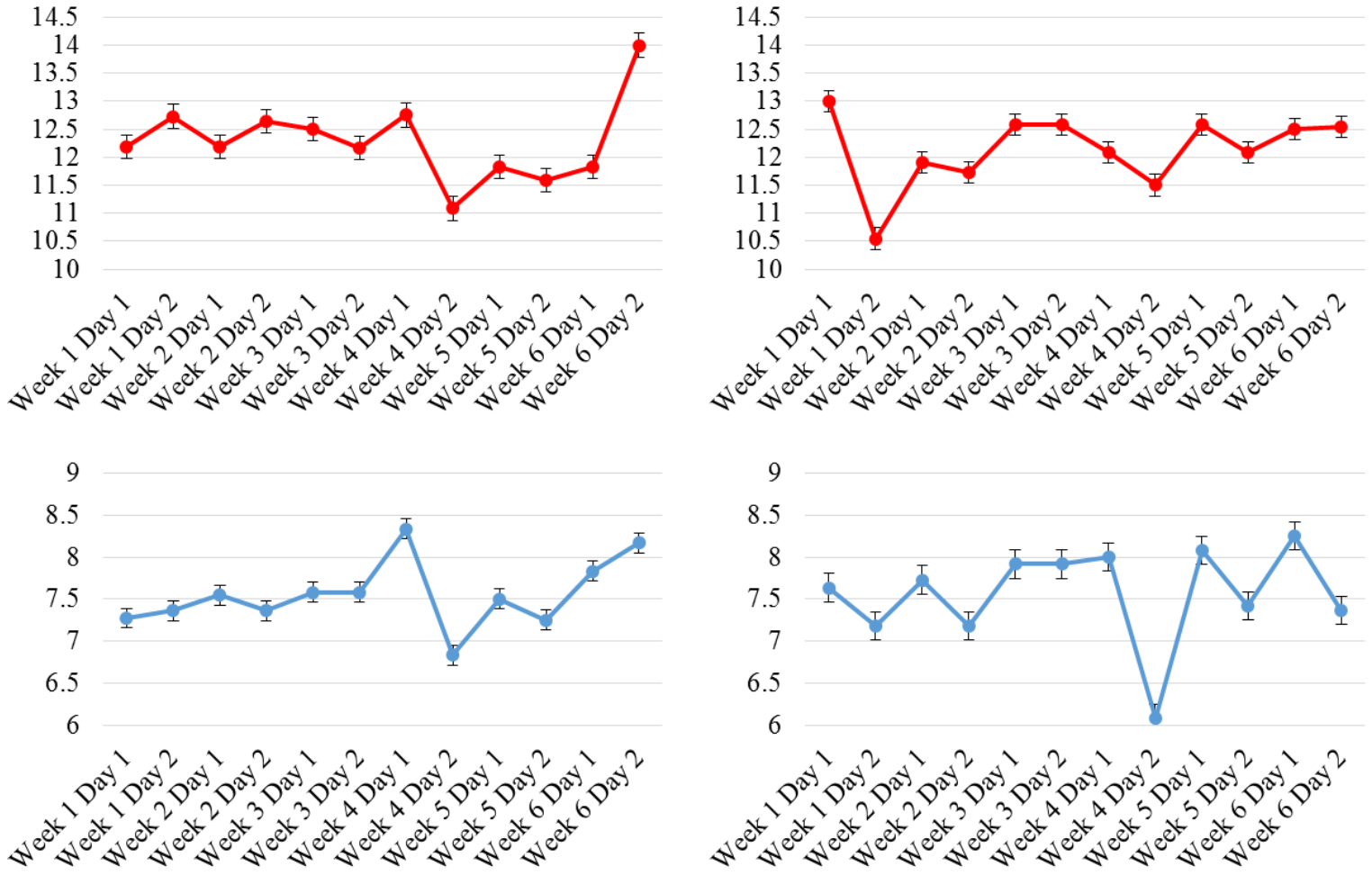
Note. Teacher-Rated Emotion Regulation for one teacher is depicted in the top left and for another teacher in the top right figure. Staff-Rated Social Functioning is depicted in the bottom left and Teacher-Rated Social Functioning is depicted in the bottom right figure.

Figure 5 Estimated Marginal Means of Teacher-Rated Emotion Regulation for two teachers as well as of Staff-Rated and of Teacher-Rated Social Functioning on the Social and Mood Ratings.



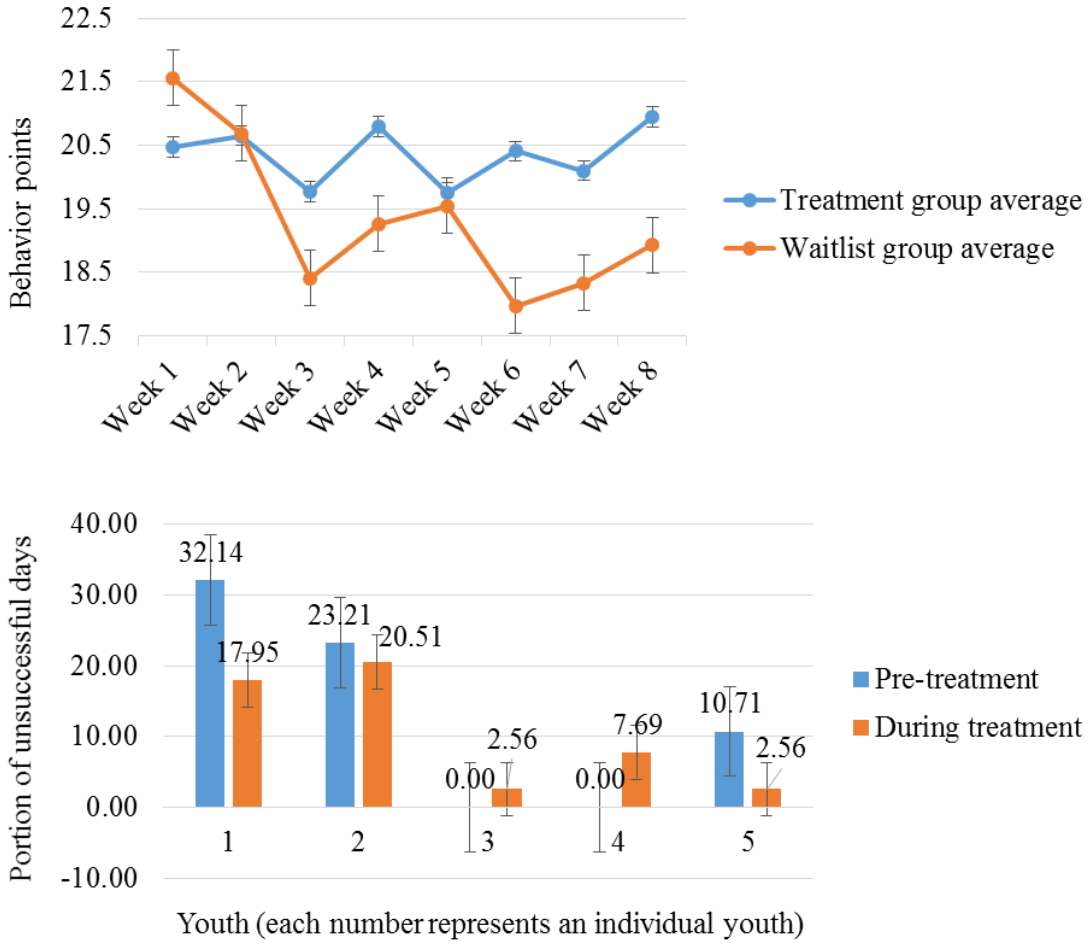
Note. Self-rated DERS Impulse is depicted in the top left, self-rated anger and verbal aggression on the BPAQ are depicted in the top right, teacher-rated IRS are depicted in the bottom left, and staff-rated property and verbal aggression on the MOAS are depicted in the bottom right figures.

Figure 6 Self-rated DERS Impulse, self-rated anger and verbal aggression on the BPAQ, teacher-rated IRS, staff-rated property and verbal aggression on the MOAS.



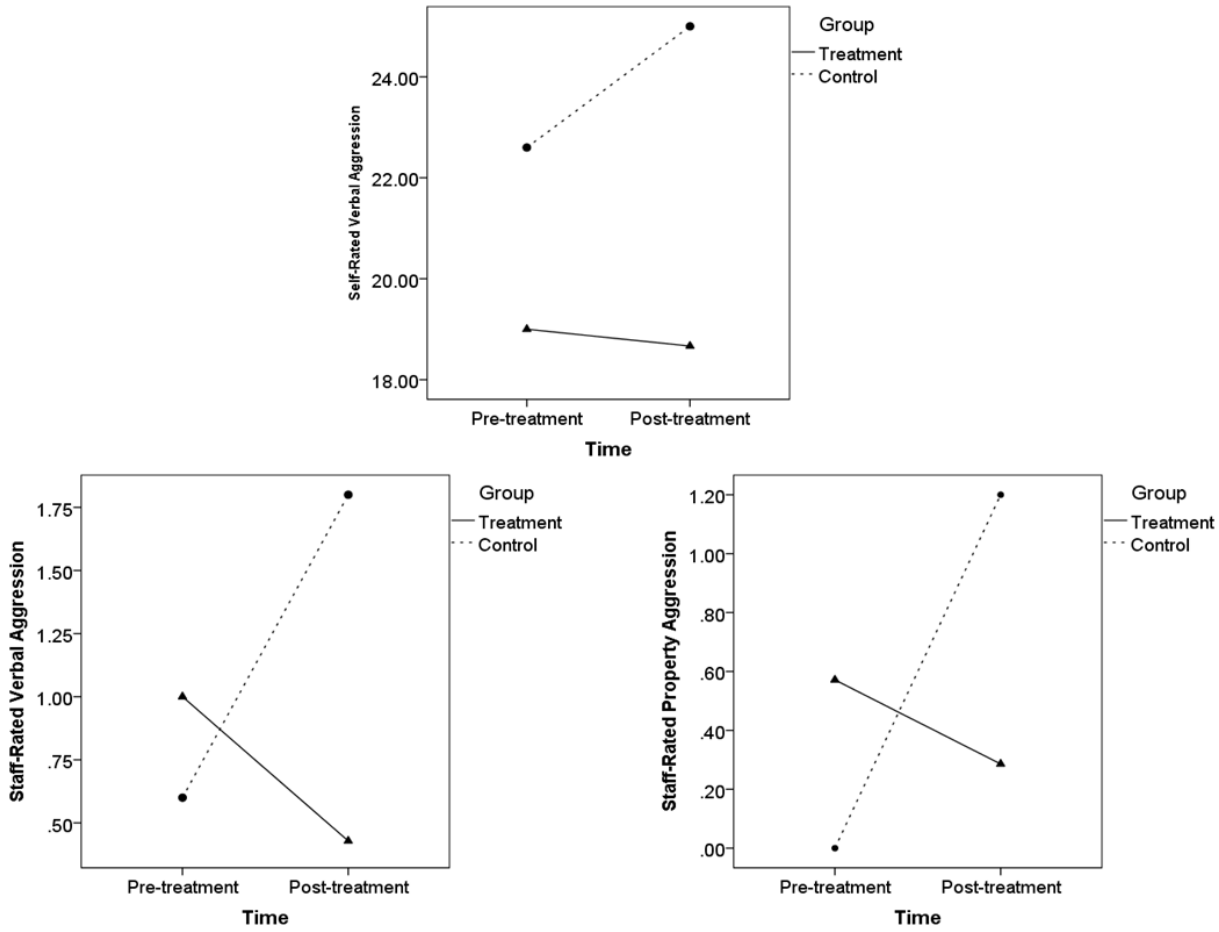
Note. Staff-rated emotion regulation is depicted in the top and social functioning is depicted in the bottom figures, with morning shift ratings depicted on the left and afternoon shift ratings depicted on the right.

Figure 7 Staff-rated emotion regulation and social functioning on the Social and Mood Ratings during the intervention.



Note. Behavior points are depicted in the top figure and portion of unsuccessful days (i.e., less than 14 behavior points per day) for each participant over time are depicted in the bottom figure.

Figure 8 Behavior points for the intervention and the waitlist group over time and portion of unsuccessful days for each participant over time.



Note. Self-rated verbal aggression is depicted in the top, staff-rated verbal aggression is depicted in the bottom left, and staff-rated property aggression is depicted in the bottom right figure.

Figure 9 Estimated marginal means of self-rated verbal aggression on the BPAQ, of staff-rated verbal aggression on the MOAS and of staff-rated property aggression on the MOAS.

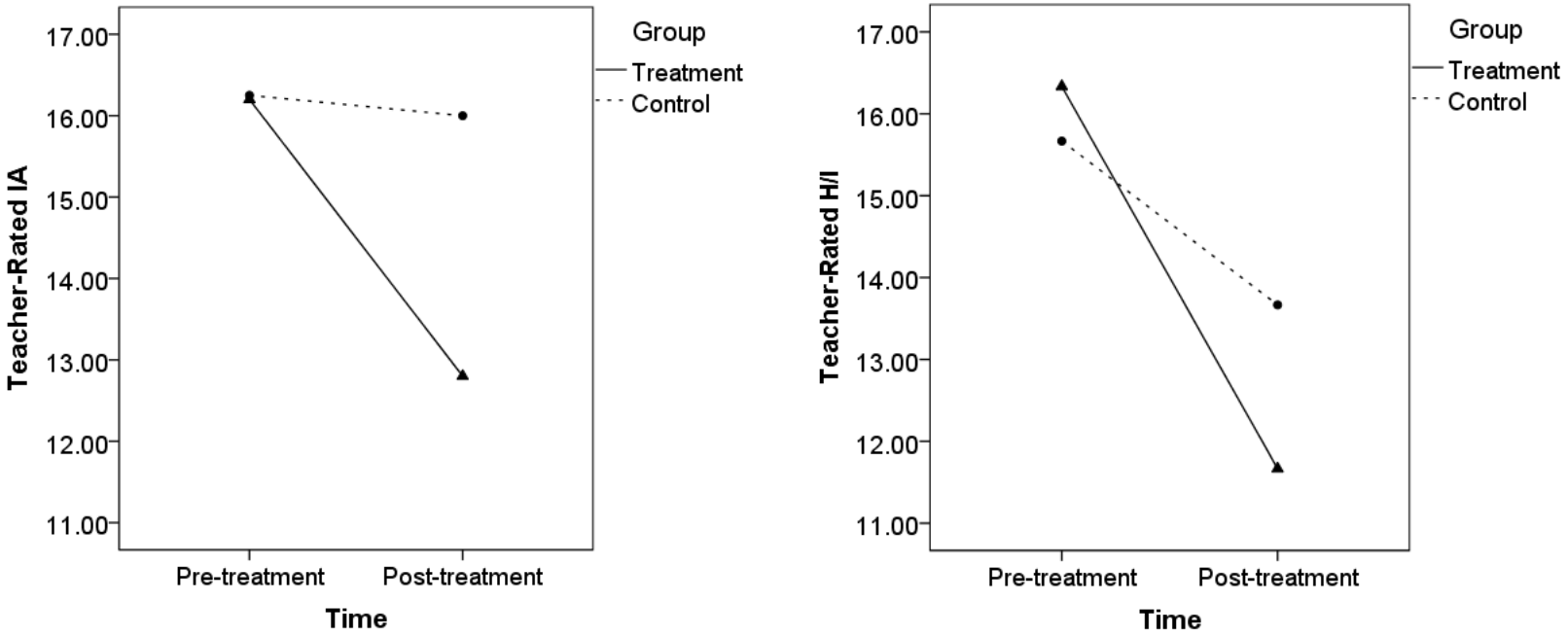


Figure 10 Estimated marginal means of teacher-rated inattention and hyperactivity/impulsivity on the Disruptive Behavior Disorders Rating Scale

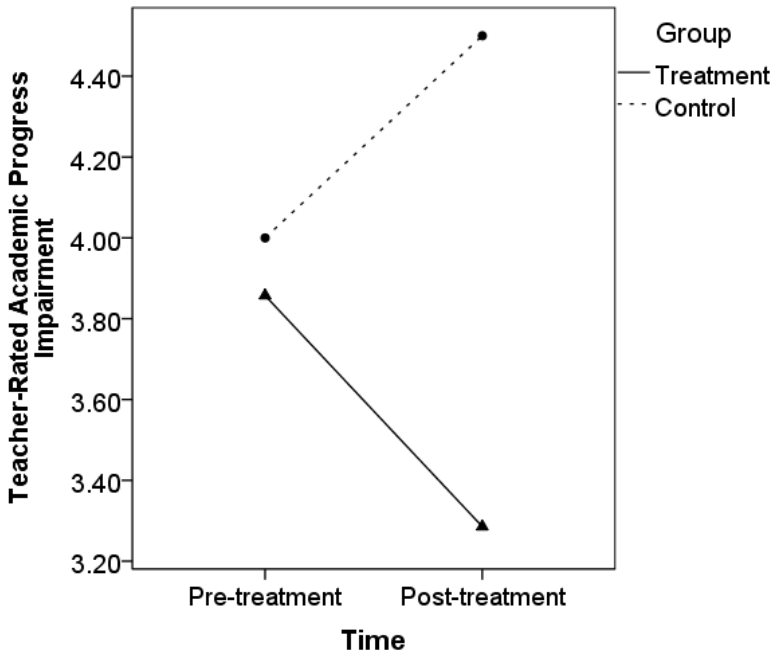


Figure 11 Estimated marginal means of teacher-rated academic progress impairment on the Impairment Rating Scale.

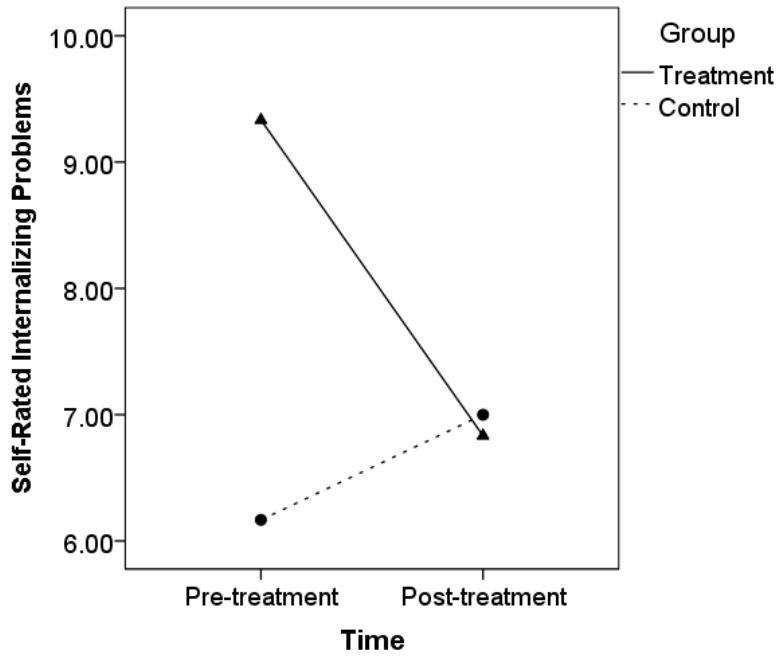


Figure 12 Estimated marginal means of self-rated internalizing problems on the Social Skills Improvement System-Rating Scale.

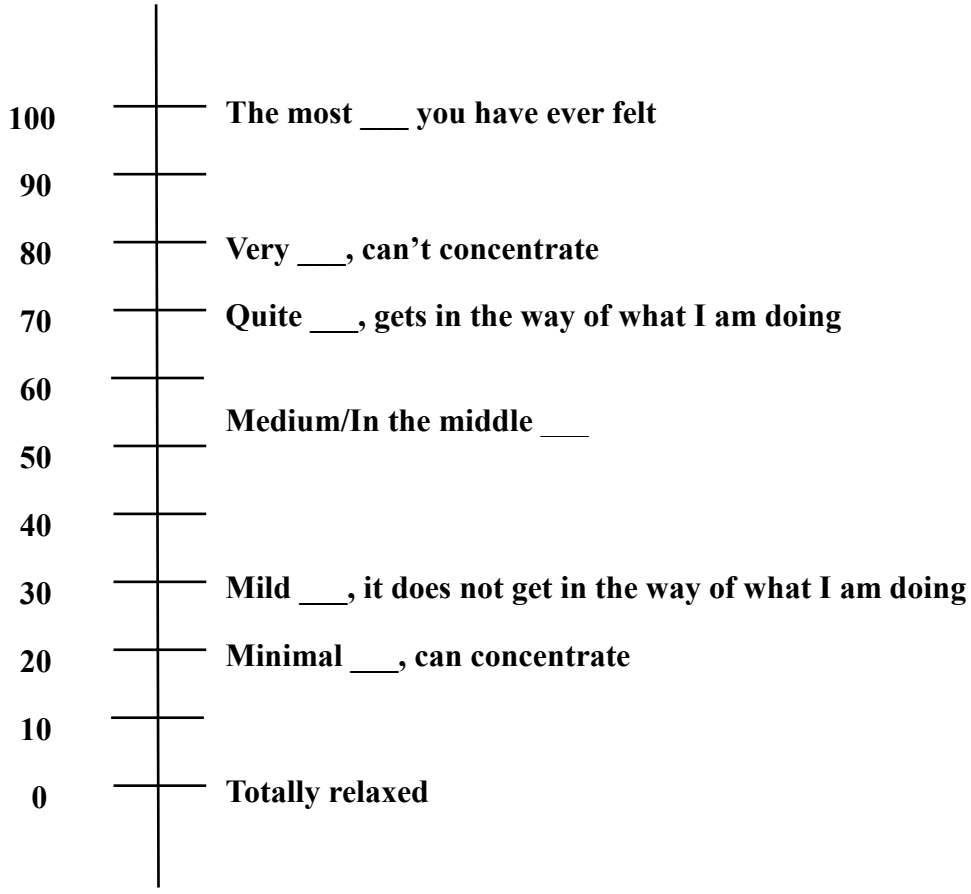
APPENDIX A: COPIES OF THE INSTRUMENTS USED IN THE STUDY

Emotions Thermometer

Rating Date _____ Rating Period _____ Participant ID _____

Pick the main emotion you have been feeling/felt during the last few minutes:
Happiness/Excitement Anger/Frustration Sadness/Disappointment

Now imagine you have an "Emotions Thermometer" to measure your emotion. On this scale, 0 is feeling perfectly relaxed and 100 is the strongest of your emotion you can imagine.
Now rate your emotion on the Emotions Thermometer



Parent / Teacher DBD Rating Scale

Child's Name: _____Form Completed by:
Grade: _____ Date of Birth: _____ Sex: _____Date Completed: _____Check the column that best describes your/this child. Please write DK next to any items for which you don't know the answer.

	Not at All	Just a Little	Pretty Much	Very Much
1. often interrupts or intrudes on others (e.g., butts into conversations or games)				
2. has run away from home overnight at least twice while living in parental or parental surrogate home (or once without returning for a lengthy period)				
3.often argues with adults				
4.often lies to obtain goods or favors or to avoid obligations (i.e., "cons" others)				
5.often initiates physical fights with other members of his or her household				
6.has been physically cruel to people				
7.often talks excessively				
8. has stolen items of nontrivial value without confronting a victim (e.g., shoplifting, but without breaking and entering; forgery)				
9.is often easily distracted by extraneous stimuli				
10. often engages in physically dangerous activities without considering possible consequences (not for the purpose of thrill-seeking), e.g., runs into street without looking				
11. often truant from school, beginning before age 13 years				
12. often fidgets with hands or feet or squirms in seat				
13. is often spiteful or vindictive				
14. often swears or uses obscene language				
15. often blames others for his or her mistakes or misbehavior				
16. has deliberately destroyed others' property (other than by fire setting)				
17. often actively defies or refuses to comply with adults' requests or rules				
18. often does not seem to listen when spoken to directly				

19. often blurts out answers before questions have been completed				
20. often initiates physical fights with others who do not live in his or her household (e.g., peers at school or in the neighborhood)				
21. often shifts from one uncompleted activity to another				
22. often has difficulty playing or engaging in leisure activities quietly				
23. often fails to give close attention to details or makes careless mistakes in schoolwork, work, or other activities				
24. is often angry and resentful				
25. often leaves seat in classroom or in other situations in which remaining seated is expected				
26. is often touchy or easily annoyed by others				
27. often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (not due to oppositional behavior or failure to understand instructions)				
28. often loses temper				
29. often has difficulty sustaining attention in tasks or play activities				
30. often has difficulty awaiting turn				
31. has forced someone into sexual activity				
32. often bullies, threatens, or intimidates others				
33. is often "on the go" or often acts as if "driven by a motor"				
34. often loses things necessary for tasks or activities (e.g., toys, school assignments, pencils, books, or tools)				
35. often runs about or climbs excessively in situations in which it is inappropriate (in adolescents or adults, may be limited to subjective feelings of restlessness)				
36. has been physically cruel to animals				
37. often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (such as schoolwork or homework)				
38. often stays out at night despite parental prohibitions, beginning before age 13 years				
39. often deliberately annoys people				
40. has stolen while confronting a victim (e.g., mugging, purse snatching, extortion, armed robbery)				

41. has deliberately engaged in fire setting with the intention of causing serious damage				
42. often has difficulty organizing tasks and activities				
43. has broken into someone else's house, building, or car				
44. is often forgetful in daily activities				
45. has used a weapon that can cause serious physical harm to others (e.g., a bat, brick, broken bottle, knife, gun)				

Ideal Self Goals Rating Card

Student's ID: Date: _____

Counselor's Name:

Ratings should indicate the degree with which the student portrayed each ideal self goal according to the scale below. After the last feedback, transfer average ratings to mastery form.

Behavior portrayed the opposite of goal					*Behavior portrayed the desired goal*		
Very Much **-3**	Some **-2**	A Little **-1**	Either Way **0**	A Little **1**	Some **2**	Very Much **3**	No Opportunity **N.O.**

Ideal Self Goals	Feedback 1	Feedback 2	AVERAGE SCORES
#1:	Self		Self
	Staff		Staff
#2:	Self		Self
	Staff		Staff

Note. Clients are asked for their self-ratings and these are recorded in the small boxes within each cell before the counselor shares his/her ratings, which are recorded in the large boxes within each cell.

DERS

Please indicate how often the following statements apply to you by writing the appropriate number from the scale below on the line beside each item:

1	2	3	4	5
Almost never	Sometimes	About half the time	Most of the time	Almost always
0-10%	11-35%	36-65%	66-90%	91-100%

_____ 1) I am clear about my feelings.

_____ 2) I pay attention to how I feel.

_____ 3) I experience my emotions as overwhelming and out of control.

_____ 4) I have no idea how I am feeling.

_____ 5) I have difficulty making sense out of my feelings.

_____ 6) I am attentive to my feelings.

_____ 7) I know exactly how I am feeling.

_____ 8) I care about what I am feeling.

_____ 9) I am confused about how I feel.

_____ 10) When I'm upset, I acknowledge my emotions.

_____ 11) When I'm upset, I become angry with myself for feeling that way.

_____ 12) When I'm upset, I become embarrassed for feeling that way.

_____ 13) When I'm upset, I have difficulty getting work done.

_____ 14) When I'm upset, I become out of control.

_____ 15) When I'm upset, I believe that I will remain that way for a long time.

_____ 16) When I'm upset, I believe that I'll end up feeling very depressed.

_____ 17) When I'm upset, I believe that my feelings are valid and important.

_____ 18) When I'm upset, I have difficulty focusing on other things.

_____ 19) When I'm upset, I feel out of control.

_____ 20) When I'm upset, I can still get things done.

_____ 21) When I'm upset, I feel ashamed with myself for feeling that way.

_____ 22) When I'm upset, I know that I can find a way to eventually feel better.

_____ 23) When I'm upset, I feel like I am weak.

_____ 24) When I'm upset, I feel like I can remain in control of my behaviors.

_____ 25) When I'm upset, I feel guilty for feeling that way.

_____ 26) When I'm upset, I have difficulty concentrating.

_____ 27) When I'm upset, I have difficulty controlling my behaviors.

_____ 28) When I'm upset, I believe that there is nothing I can do to make myself feel better.

_____ 29) When I'm upset, I become irritated with myself for feeling that way.

_____ 30) When I'm upset, I start to feel very bad about myself.

_____ 31) When I'm upset, I believe that wallowing in it is all I can do.

_____ 32) When I'm upset, I lose control over my behaviors.

_____ 33) When I'm upset, I have difficulty thinking about anything else.

_____ 34) When I'm upset, I take time to figure out what I'm really feeling.

_____ 35) When I'm upset, it takes me a long time to feel better.

_____ 36) When I'm upset, my emotions feel overwhelming.

ERICA

Below are a number of statements. Please read each statement and then circle the choice that seems **most true for you**. Do not spend too much time on any one item. Remember, this is not a test. There are no right or wrong answers. We really want to know what you think.

1.	I am a happy person	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
2.	When adults are friendly to me, I am friendly to them	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
3.	I handle it well when things change or I have to try something new	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
4.	When I get upset, I can get over it quickly	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
5.	*When things don't go my way I get upset easily*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
6.	When other kids are friendly to me, I am friendly to them	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
7.	*I have angry outbursts*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
8.	*I enjoy seeing others hurt or upset*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
9.	*I can be disruptive at the wrong times*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
10.	*I get angry when adults tell me what I can and cannot do*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
11.	*I am a sad person*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
12.	*I have trouble waiting for something I want*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
13.	*I am quiet and shy, and I don't show my feelings*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
14.	*I do things without thinking about them first*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
15.	When others are upset, I become sad or concerned for them	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree
16.	*I annoy others by not minding my own business*	Strongly Disagree	Disagree	Half and Half	Agree	Strongly Agree

Buss-Perry Scale

Please rate each of the following items in terms of how characteristic they are of you. Use the following scale for answering these items.

1 2 3 4 5 6 7
extremely uncharacteristic of me **extremely characteristic of me**

1) Once in a while I can't control the urge to strike another person.
2) Given enough provocation, I may hit another person.
3) If somebody hits me, I hit back.
4) I get into fights a little more than the average person.
5) If I have to resort to violence to protect my rights, I will.
6) There are people who pushed me so far that we came to blows.
7) I can think of no good reason for ever hitting a person.
8) I have threatened people I know.
9) I have become so mad that I have broken things.
10) I tell my friends openly when I disagree with them.
11) I often find myself disagreeing with people.
12) When people annoy me, I may tell them what I think of them.
13) I can't help getting into arguments when people disagree with me.
14) My friends say that I'm somewhat argumentative.
15) I flare up quickly but get over it quickly.
16) When frustrated, I let my irritation show.
17) I sometimes feel like a powder keg ready to explode.
18) I am an even-tempered person.
19) Some of my friends think I'm a hothead.
20) Sometimes I fly off the handle for no good reason.
21) I have trouble controlling my temper.
22) I am sometimes eaten up with jealousy.
23) At times I feel I have gotten a raw deal out of life.
24) Other people always seem to get the breaks.
25) I wonder why sometimes I feel so bitter about things.
26) I know that "friends" talk about me behind my back.
27) I am suspicious of overly friendly strangers.
28) I sometimes feel that people are laughing at me behind me back.
29) When people are especially nice, I wonder what they want.

THE MODIFIED OVERT AGGRESSION SCALE (MOAS)*

Patient

Rater

INSTRUCTIONS

Rate the patient's aggressive behavior over the past week. Select as many items as are appropriate. Refer to the pocket guide for the full measure.

SCORING

1. Add items in each category
2. In scoring summary, multiply sum by weight and add weighted sums for total weighted score. Use this score to track changes in level of aggression over time.

Verbal aggression

0 No verbal Aggression
1 Shouts angrily, curses mildly, or makes personal insults
2 Curses viciously, is severely insulting, has temper outbursts
3 Impulsively threatens violence toward others or self
4 Threatens violence toward others or self repeatedly or deliberately

Aggression against Property

0 No aggression against property
1 Slams door, rips clothing, urinates on floor
2 Throws objects down, kicks furniture, defaces walls
3 Breaks objects, smashes windows
4 Sets fires, throws objects dangerously

Autoaggression

0 No autoaggression
1 Picks or scratches skin, pulls hair out, hits self (without injury)
2 Bangs head, hits fists into walls, throws self onto floor
3 Inflicts minor cuts, bruises, burns, or welts on self
4 Inflicts major injury on self or makes a suicide attempt

Physical Aggression

0 No physical aggression
1 Makes menacing gestures, swings at people, grabs at clothing
2 Strikes, pushes, scratches, pulls hair of others (without injury)
3 Attacks others, causing mild injury (bruises, sprain, welts, etc.)
4 Attacks others, causing serious injury

CATEGORY	SUM SCORE	WEIGHTS	WEIGHTED SUM
Verbal Aggression		x 1	
Aggression against Property		x 2	
Autoaggression		x 3	
Physical Aggression		x 4	
Total Weighted Score			

*Modified from Kay SR, Wolkenfelf F, Murrill LM (1988), Profiles of aggression among psychiatric patients: I. nature and prevalence. *Journal of Nervous and Mental Disease 176*:539-546

Social and Mood Ratings

Week#: ___ (Week of _____) Child name (in pencil) _____

Day 1/Time 1 (day shift)

Staff Name _____

Social and Mood Ratings – Staff Form

Please, read each item carefully and circle the number that best describes this child's behavior for the past 2-4 days, to the best of your knowledge.

	Always/ Almost Always		Sometimes		Never/ Almost Never
Gets along with staff.	5	4	3	2	1
Gets along with peers.	5	4	3	2	1
Is overly animated or excited.	5	4	3	2	1
Is irritable, easily frustrated, short-tempered, or moody.	5	4	3	2	1
Is withdrawn, indifferent, disinterested, gloomy or mopes around.	5	4	3	2	1
Handles changes well, is flexible and easygoing.	5	4	3	2	1

*Items in italics are reverse-scored.

***Office use only:

Assessment Date _____ Assessment Time _____

Participant ID _____

Impairment Rating Scale
TEACHER-- Narrative

ID# _____ Form Completed by: _____

Date Completed: _____

Instructions: In the space below, please describe what you see as this child's primary problems. Also, please describe how this child's problems have affected the following areas and complete the rating at the end of each: (1) his or her relationships with other children, (2) your relationship with him or her, (3) his or her academic progress (4) your classroom in general, and (5) his or her self-esteem. Continue on a separate sheet if necessary. **For the ratings, please mark an "X" on the lines at the points that you believe reflect the impact of the child's problems in this area and whether he or she needs treatment or special services for the problems. PLEASE COMPLETE BOTH SIDES OF THIS FORM.**

(1) How this child's problems affect his or her relationship with other children

No Problem | Extreme Problem
Definitely does not need treatment or special services | Definitely needs treatment or special services

(2) Regardless of whether this child is popular or unpopular with peers, does he or she have a special, close "best friend" that he or she has kept for more than a few months? (Please circle)

YES NO

(3) How this child's problems affect his or her relationship with you, the teacher

No Problem | Extreme Problem
Definitely does not need treatment or special services | Definitely needs treatment or special services

PLEASE COMPLETE THE REVERSE SIDE OF THIS FORM

(4) How this child's problems affect his or her academic progress

No Problem
Definitely does not need treatment or special services

Extreme Problem
Definitely needs treatment or special services

(5) How this child's problems affect your classroom in general

No Problem
Definitely does not need treatment or special services

Extreme Problem
Definitely needs treatment or special services

(6) How this child's problems affect his or her self-esteem

No Problem
Definitely does not need treatment or special services

Extreme Problem
Definitely needs treatment or special services

(7) Please mark an "X" on the following line at the point that you believe reflects the overall severity of this child's problem in functioning and overall need for treatment.

No Problem
Definitely does not need treatment or special services

Extreme Problem
Definitely needs treatment or special services

Satisfaction Questionnaire

Please help us improve our program by answering some questions about the services you have received. We are interested in your honest opinion, whether they are positive or negative.
Please answer all of the questions we also welcome your comments and suggestions. Thank you very much, we really appreciate your help.

CIRCLE YOUR ANSWER

1. How would you rate the quality of service you have received?

4	3	2	1
Excellent	*Good*	*Fair*	*Poor*

2. Did you get the kind of service you wanted?

1	2	3	4
No, definitely not	*No, not really*	*Yes, generally*	*Yes, definitely*

3. To what extent has our program met your needs?

4	3	2	1
Almost all of my needs have been met	*Most of my needs have been met*	*Only a few of my needs have been met*	*None of my needs have been met*

4. If a friend were in need of similar help, would you recommend our program to him or her?

1	2	3	4
No, definitely not	*No, I don't think so*	*Yes, I think so*	*Yes, definitely*

5. How satisfied are you with the amount of help you have received?

1	2	3	4
Quite dissatisfied	*Indifferent or mildly dissatisfied*	*Mostly satisfied*	*Very satisfied*

6. Have the services you received helped you to deal more effectively with your problems?

4	3	2	1
Yes, they helped a great deal	*Yes, they helped somewhat*	*No, they really didn't help*	*No, they seemed to make things worse*

7. You came to our program with certain problems. How are those problems now?

1	2	3	4
Worse or much worse	*No change*	*Somewhat better*	*A great deal better*

8. In an overall, general sense, how satisfied are you with the service you have received

4	3	2	1
Very satisfied	*Mostly satisfied*	*Indifferent or mildly dissatisfied*	*Quite dissatisfied*

9. If you were to seek help again, would you come back to our program?

1	2	3	4
No, definitely not	*No, I don't think so*	*Yes, I think so*	*Yes, definitely*

Preferred Treatment Elements Questionnaire

You just received an intervention, the TEAM group, from our staff. Now, we would like to ask you to share your opinion about this group with us.

1. First, we would like for you to tell us *what you liked most* about the TEAM group.
2. Then, we would like for you to tell us what parts of the TEAM group you *found to be most helpful to you*.

For some people, these may be the same but for others they may be different. There are no right or wrong answers. Please, just tell us how you feel/what you think about the TEAM group.

Preferred Elements

Please, rank the following treatment elements in light of *what you liked most*, by connecting the number on the left with the treatment elements on the right: "1" with the one you liked the best, "2" with the one you liked second best and "3" with the one you liked least.

1	Psychoeducation/Phase I (Session #1) – This was the first session when we discussed things like Ideal Self, Real Self, Emotions and Emotion Regulation. We also practiced Mindfulness Meditation and created Ideal Self Goals.
2	Practice Activities/Phase II (Sessions #2 through 13) – During these sessions we played a number of group games. You also rated your progress on your Ideal Self Goals and a counselor rated you as well. We also used the Emotions Thermometer to rate the intensity of your emotions.
3	At-home mindfulness meditation practice – This is when you listened to a mindfulness mediation relaxation exercise on your MP3 player.

Please, explain why you ranked Psychoeducation for **Preferred Elements** as you did:

_____

_____

Please, explain why you ranked Practice Activities for **Preferred Elements** as you did:

_____

_____

Explain why you ranked At-home mindfulness meditation for **Preferred Elements** as you did:

_____

_____

Beneficial/Helpful Elements

Please, rank the following treatment elements in light of *what you found most helpful* to you, by connecting the number on the left with the treatment elements on the right: "1" with the one you found most helpful, "2" with the one you found second most helpful and "3" with the one you found least helpful.

1	Psychoeducation/Phase I (Session #1) – This was the first session when we discussed things like Ideal Self, Real Self, Emotions and Emotion Regulation. We also practiced Mindfulness Meditation and created Ideal Self Goals.
2	Practice Activities/Phase II (Sessions #2 through 13) – During these sessions we played a number of group games. You also rated your progress on your Ideal Self Goals and a counselor rated you as well. We also used the Emotions Thermometer to rate the intensity of your emotions.
3	At-home mindfulness meditation practice – This is when you listened to a mindfulness mediation relaxation exercise on your MP3 player.

Please, explain why you ranked Psychoeducation for **Beneficial/Helpful** as you did:

_____

_____

Please, explain why you ranked Practice Activities for **Beneficial/Helpful** as you did:

_____

_____

Explain why you ranked At-home mindfulness meditation for **Beneficial/Helpful** as you did:

_____

_____

APPENDIX B: DETAILED REVIEW OF THE PSYCHOMETRIC PROPERTIES OF THE INSTRUMENTS USED IN THE STUDY

Buss–Perry Aggression Questionnaire (BPAQ; Buss & Perry, 1992). The BPAQ is a 29-item, self-rated five-point Likert scale designed to assess four dimensions, including physical aggression, verbal aggression, anger, and hostility (Buss & Perry, 1992). Internal consistency for the four subscales and total score range from .72 for Verbal Aggression to .89 for the Total score. Retest reliability over 9 weeks ranged from .72 for the Anger subscale to .80 for the Physical Aggression subscale and Total score. Prior findings indicate that the BPAQ is sensitive to treatment effects (Shelton et al., 2011).

Children's Interview for Psychiatric Syndromes (ChIPS*;* Weller, Weller, Fristad, Rooney, & Schechter, 2000). The ChIPS is a semi-structured diagnostic interview of 20 DSM-IV-TR (APA, 2000) mental health disorders in youth 6-18 years. The ChIPS has adequate psychometric properties (Weller, Weller, Fristad, Rooney, & Schecter, 2000). Specifically, there is moderate agreement among the ChIPS and clinician diagnoses (α = 0.49). Further, the average sensitivity is 87% and specificity is 76% (Fristad et al., 1998).

Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). The DERS is a 36-item self-report measure of ED. Items are rated on a 5-point Likert-type scale (1 = Almost Never, 5 = Almost Always), with higher scores indicating greater difficulty with emotion regulation. Subscales of the DERS are as follows: Nonacceptance of Emotional Responses (Nonacceptance), Difficulties Engaging in Goal-Directed Behavior (Goals), Impulse Control Difficulties (Impulse), Lack of Emotional Awareness

(Awareness), Limited Access to Emotion Regulation Strategies (Strategies), and Lack of Emotional Clarity (Clarity). Scores for the total DERS range from 36 to 180. The DERS demonstrated good internal consistency (α = .924; Bunford et al., 2014) as well as good test–retest reliability, and adequate construct and predictive validity in multiple adolescent samples (Adrian et al., 2009; Vasilev et al., 2009; Weinberg & Klonsky, 2009). In addition, the DERS exhibited robust correlations with psychological problems reflecting ED (Weinberg & Klonsky, 2009) and physiological measures of ED (Vasilev et al., 2009). Prior findings indicate that the DERS is sensitive to treatment effects (e.g., Axelrod et al., 2011).

Disruptive Behavior Disorders Rating Scale – Parent and - Teacher Report (DBD; Pelham et al., 1992, see Appendix). The 45-item DBD scale targets symptom criteria established for the three disruptive behavior disorders, ADHD, ODD and CD. Questions on the DBD correspond to DSM symptoms of the three disorders.

Not many studies have examined the psychometrics and factor structure of the DBD – Parent, particularly for young adolescents. Results using DSM-III-R criteria in one study with 5- to 15-year-old boys indicated that two factors best represented ADHD items when modeled with ODD and CD items, but three factors best represented ADHD items when modeled without ODD and CD items (Pillow, Pelham, Hoza, Molina, & Stultz; 1998). In a more recent study, Zuddas et al., found support for two ADHD factors—H/I and inattention (Zuddas et al., 2006) using DSM-IV criteria. More recently, Van Eck, Finney, and Evans (2010) examined the ADHD items only of the scale. A two-factor model of IA and H/I provided the best (i.e., adequate) fit to item responses,

consistent with findings from previous studies of teacher report (Pelham et al., 1992). Internal consistency was adequate for both H/I and IA. Although the IA factor explained an acceptable amount of variance in inattention items (54%), the H/I factor explained only 45% of variance in hyperactivity/impulsivity items. Finally, a moderate correlation (r = .55) was obtained between the two factors, suggesting distinct but related constructs.

Using ratings from secondary school teachers of 9th through 12th graders, reliability estimates of the DBD – Teacher were found to be adequate and above .90 for the three inattention (IA), impulsivity-hyperactivity (H/I), and ODD subscales (Evans, et al., 2013). When compared to other measures of closely related constructs such as the Teacher Report Form (a measure of teachers' reports of children's academic performance, adaptive functioning, and behavioral/emotional problems; TRF; Achenbach & Edelbrock, 1986) and the Iowa/Abbreviated Conners Teacher Rating Scale (a measure of inattention/overactivity and aggression/defiance; IOWA; Goyette et al., 1978), correlations between the DBD – Teacher and the dimensional subscales of the TRF and the IOWA were all statistically significant and medium to large (ranged from .55 to .83) (Molina, Pelham, Blumenthal, & Galiszevki, 1998). Prior findings indicate that the DBD – Parent and Teacher are sensitive to treatment effects (e.g., Evans et al., 2016).

Emotion Regulation Index for Children and Adolescents (ERICA; MacDermott et al., 2010). The ERICA is a 16-item self-report measure of emotion regulation in children and adolescents. Items are rated on a 5-point Likert-type scale (1 = Strongly Disagree to 5 = Strongly Agree), with higher scores indicating better regulation of emotions. Subscales of the ERICA are as follows: Emotional Control (Emotional Control), Emotional Self-

Awareness (Self-Awareness), and Situational Responsiveness (Situational Responsiveness). Scores for the total ERICA range from 16 to 80. The ERICA has demonstrated good internal consistency (α = .831), Bunford et al., 2014), good test-retest reliability, and good convergent and construct validity (MacDermott et al., 2010).

Heart-Rate Variability data acquisition and processing. Heart rate variability (i.e., the oscillation in the interval between consecutive heartbeats), was chosen as a quantitative marker of autonomic nervous system activity as it represents one of the best such markers (Malik, 1996). HRV data was collected during a five-minute resting baseline (prior to youth working on the frustration task), during a five-minute frustration task, and during a five-minute recovery (following youth working on the frustration task). HRV was recorded using a Polar Pro Trainer RS800 system (Polar Electro, Kempele, Finland), comprised of a wristwatch and a bio-harness. Youth wore the bio-harness on their upper torsos with the center of the harness located over the xiphisternum. This system is used to collect data on and analyze HRV in a non-invasive manner while estimating the function of the sympathetic and parasympathetic autonomic nervous systems. The system fulfills the requirements of the Heart Rate Variability Standards (Task Force of the European Society of Cardiology and Task Force of the European Society of Cardiology, 1996). Data was recorded with Polar Pro-Trainer 5 software (PPT 5) (Polar Electro, Kempele, Finland). Data on variability in the time interval between heartbeats was extracted and corrected for artifacts using Kubios HRV Analysis Software 2.0 (Tarvainen, Niskanen, Lipponen, Ranta-Aho, & Karjalainen, 2014).

Kubios HRV 2.0 was used to calculate the root mean square of successive

differences (RMSSD) between adjacent time intervals (measured in msec) between heartbeats, to index HRV. Higher values indicate less ED. RMSSD is a stable (Li et al., 2009) and valid (Thayer & Sternberg, 2010) time-domain measure of HRV.

Impairment Rating Scale – Teacher Report (IRS; Fabiano et al., 2006). The IRS is a teacher-report measure of areas of functioning in which children with ADHD are typically impaired (Fabiano et al., 2006). The IRS contains six items (relationship with peers, relationship with teacher, academic progress, self-esteem, influence on classroom functioning, and overall impairment). Teachers rate items on a visual analogue scale by placing an "X" on a line that represents a continuum of impairment and ranges from 0 ("no problem/definitely does not need treatment or special services") to 6 ("extreme problem/definitely needs treatment or special services"). As each item represents a unique domain of impairment, scores range from 0 to 6. The IRS demonstrated excellent temporal stability and convergent and discriminant validity (Fabiano et al., 2006) as well as effectiveness in discriminating between children whose parent and teacher ratings of ADHD symptoms are consistent with a diagnosis of ADHD and children whose ratings of symptoms is below the DSM-IV threshold (Fabiano et al., 2006). The authors of prior studies did not find that the IRS is sensitive to treatment effects (e.g., Evans et al., 2016).

Modified Overt Aggression Scale (MOAS; Coccaro, Harvey, Kupsaw-Lawrence, Herbert, & Bernstein, 1991). The MOAS is a 25-item measure of staff-rated severity, type, and frequency of aggressive behavior. Items are weighted by severity and frequency. Subtypes include: verbal aggression, physical aggression against objects, and physical aggression against self and physical aggression against others (Coccaro, Harvey,

Kupsaw-Lawrence, Herbert, & Bernstein, 1991). Inter-rater reliability has been demonstrated to be .91 for ratings by two clinical raters for MOAS Aggression and Irritability. Test–retest reliability within a 1 to 2-week period has been shown to have an intraclass correlation for aggression on Time 1 and Time 2 of .46 and .54, respectively (Suris et al., 2005). The authors of prior studies did not find that the MOAS is sensitive to treatment effects (e.g., Shelton et al., 2011).

PMCJF behavior points. Youth behavior and participation in PMCJF activities is evaluated on a point-system as part of regular facility functions. The structure of the point system is as follows: All youth begin their day having 20 points and can lose points given misbehavior or insufficient participation in facility activities. During the first half of each day, youth have three points for school, two points for daily living, two points for following facility rules and three points for not having any incident reports (incident reports are written by facility staff in the case of serious aggression or rule-violations). During the second half of the day, youth have three points for following rules in groups, two points for daily living, two points for following facility rules and three points for not having any incident reports. Should a youth not lose any points during either the first or the second half of the day, he retains all 20 points for that day. Youth who retain all 20 points are considered to have "made their day" which, in turn, makes them eligible for two additional points, totaling to 22 points (youth can either exchange some of their 20 points for privileges or items from the commissary – in which case they do not earn two extra points, or they can keep their 20 points – in which case they earn two extra points). In addition to being able to exchange points for privileges or items from the commissary,

youths' behavior points are used as an index of youths' compliance with, participation in, and progression with the program that is then reported to their probation officers and ultimately the courts and thus taken into account for decisions about youths' release. No psychometric data are available.

Preferred Treatment Elements Questionnaire (PTEQ; Bunford & Evans, unpublished). The PTEQ was used to measure youth perception of treatment elements with regard to the degree to which those are likeable and beneficial/helpful. The PTEQ was created for this study. No psychometric data are available.

Satisfaction Questionnaire (Bunford & Evans, unpublished). A nine-item Satisfaction Questionnaire was used to measure treatment satisfaction. The Satisfaction Questionnaire was adapted from Nguyen et al (1983) and created for this study. Higher values indicate better satisfaction. No psychometric data are available.

Social and Mood Ratings (SMR; Bunford & Evans, unpublished). The six-item SMR is a measure of staff- and teacher-rated ED and social impairment. The SMR was created for this study. Higher values on both the emotion regulation and the social functioning items indicate better emotion regulation and better social functioning, respectively. No psychometric data are available.

Social Skills Improvement System-RS – Self-Report (SSIS-RS; Elliott & Gresham, 2008, see Appendix). The SSIS-RS – Self-Report is a measure of two domains of social functioning: social skills and problem behaviors. The social skills domain is comprised of 46 items and the problem behaviors domain is comprised of 29 items, rated on a 4-point scale (0 = Never to 3 = Almost Always). Scores on the social skills subscale range from 0

to 138 and on the problem behaviors ranges from 0 to 87. Norms and standardization sample information reveal acceptable psychometric properties: median scale reliabilities of the Social Skills and Problem Behavior scales in the mid- to upper .90s; median subscale reliabilities near .80; test-retest indices for Total Social Skills .81; test-retest indices for Total Problem Behavior .77 (Gresham, Elliott, Vance, & Cook, 2011). The authors of prior studies did not find that the SSIS-RS is sensitive to treatment effects (e.g., Evans et al., 2016).

Wechsler Abbreviated Scale of Intelligence – II (WASI-II; Wechsler, 1999)**.** The WASI was developed to provide a short and reliable estimate of cognitive ability of individuals six to 89 years (Wechsler, 1991). The two-subtest form (FSIQ-2) including the vocabulary and matrix reasoning subtests provides an estimate of general cognitive ability. Adequate reliability has been demonstrated for the two-subtest form (Wechsler, 1999). The 2-subtest WASI FSIQ is highly correlated with the Wechsler Intelligence Scale for Children-Fourth Edition FSIQ (r = .83; Homack & Reynolds, 2007). The FSIQ-2 will be administered to adolescents in this study.

Weekly Ideal Self Goal ratings. Self- and clinician ratings of Ideal Self Goal mastery occur twice during each ISG Phase II session. Youth rate their own behavior and clinicians observe youth and rate youths' behavior during group activities. Ratings are based both on youths' behavior and others' reactions to such behavior. Higher values indicate better social functioning (on a scale of "- 3" to "+ 3", "- 3" indicates behaviors that likely result in an impression by others that is incongruent with the youth's Ideal Self Goal and "+ 3" indicates behaviors that likely result in an impression by others that is

congruent with the youth's Ideal Self Goal. "0" indicates a neutral rating and reflects "no evidence either way"). No psychometric data are available.

Thesis and Dissertation Services