2022

THE ROLE OF AFFECTIVE HETEROGENEITY ON TREATMENT EFFECTS FOR YOUTH WITH CONDUCT PROBLEMS

Pevitr Singh Bansal
University of Kentucky, pevitrb@gmail.com
Author ORCID Identifier:
https://orcid.org/0000-0001-7340-4215
Digital Object Identifier: https://doi.org/10.13023/etd.2022.219

Pevitr Singh Bansal, Student

Dr. Michelle M. Martel, Major Professor

Dr. Mark T. Fillmore, Director of Graduate Studies

THE ROLE OF AFFECTIVE HETEROGENEITY ON TREATMENT EFFECTS FOR YOUTH WITH CONDUCT PROBLEMS

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Pevitr Singh Bansal
Lexington, Kentucky
Director: Dr. Michelle Marie Martel, Professor of Psychology
Lexington, Kentucky
2022

https://orcid.org/0000-0001-7340-4215

ABSTRACT OF DISSERTATION

THE ROLE OF AFFECTIVE HETEROGENEITY ON TREATMENT EFFECTS FOR YOUTH WITH CONDUCT PROBLEMS

Conduct problems (CP) are a class of disruptive and aggressive behaviors (e.g., aggression, vandalism) comprised of both oppositional defiant and conduct disorder. CP are highly heterogenous and one vital factor that parses out this heterogeneity is affect, specifically the affective traits of irritability (IRR) and limited prosocial emotions (LPE). The current study examined how IRR and LPE predict distinct aspects of treatment efficacy including (1) treatment response (i.e., magnitude of change from week 1 to week 5); (2) trajectories (i.e., shape of symptom change); and (3) time-out (i.e., behavioral and emotional reactions to time-out). Participants were 49 youth aged 7-12 years (71.4% male, 77.6% meeting CP criteria) who participated in an intensive behavioral summer treatment program. Outcomes were observational ratings of CP and parent-ratings of impairment. Pre-treatment IRR predicted reduced CP at week 5. Neither IRR nor LPE exerted any significant effect upon impairment treatment response, nor trajectories of change, nor behavioral or emotional reactions to time-out. IRR may be particularly important to consider in treating CP in youth. Future research should employ measurement of the distinct underlying treatment mechanisms (e.g., emotion dysregulation, empathy) to effectively target the mechanisms contributing to aggression and impairment in youth with different affective profiles.

KEYWORDS: Conduct Problems, Irritability, Limited Prosocial Emotions, Behavior Therapy, Treatment Response, Time-Out

Pevitr Singh Bansal

(Name of Student)

05/19/2022

Date

THE ROLE OF AFFECTIVE HETEROGENEITY
ON TREATMENT EFFECTS FOR YOUTH WITH
CONDUCT PROBLEMS

By
Pevitr Singh Bansal

Dr. Michelle M. Martel
Director of Dissertation

Dr. Mark T. Fillmore
Director of Graduate Studies

05/19/2022
Date

DEDICATION

To those who have been by my side throughout this journey – my family.

ACKNOWLEDGMENTS

Science is not the result of the effort from one single person. First, I cannot fully express my appreciation and gratitude for my dissertation chair and research advisor, Dr. Michelle Martel. She has provided me endless support throughout my graduate training and has challenged me to grow as a clinical scientist. I am grateful to have had a mentor who, over these past five years, I could learn from and laugh with (and even cry to). I am also greatly appreciative of the support from my dissertation committee: Dr. Kathleen Aspiranti, Dr. Christal Badour, Dr. Joseph Hammer, Dr. Peggy Keller, and Dr. Shannon Sauer-Zavala. Their input and feedback significantly strengthened the quality of my project. Lastly, I thank all of my friends at the University of Kentucky, including my RISK lab mates and colleagues throughout the clinical program. Their friendship, support, and knowledge mean the world to me.

Beyond the University of Kentucky, I want to thank the Penn State Hershey psychiatry faculty and the 2018 and 2019 Summer Treatment Program staff who assisted in collecting these data. Further, I am deeply indebted to the ongoing support from Dr. Dan Waschbusch. Ten years ago, I was an undergraduate counselor at the Summer Treatment Program under Dr. Waschbusch's supervision. A decade later, I am honored to continue our work together as colleagues and look forward to our future collaboration.

None of this happens without my family and friends. From celebrating my successes, to providing me comfort during the lows of graduate school and life, I thank each and every one of you. There are far too many people to name, and I am blessed to have so many amazing and supportive people in my life. Thanks for sticking with me throughout this roller coaster. From the bottom of my heart, thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Conduct problems (CP) refer to a range of disruptive and antisocial behaviors that violate the rights of others and/or established societal norms including defiance towards authority figures, vandalism, and physical aggression/cruelty towards people and/or animals (Frick et al., 2014b). As per the Fifth Edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013), CP include oppositional defiant disorder and conduct disorder (Frick & Morris, 2004). The lifetime prevalence of CP is estimated to be between seven and 10% (Ghandour et al., 2019; Nock et al., 2006, 2007), making it one of the most common referral reasons to mental health clinics (Loeber et al., 2000). Youth with CP experience substantial impairments across several life domains including interpersonal, academic, and psychological functioning (e.g., Waschbusch, 2002). Further, these youth also pose a significant financial burden to caregivers as the cost to care for a child with CP is approximately $70,000 greater than that for a typically developing child (Foster et al., 2005). In the absence of treatment, youth with CP are at a significantly elevated risk to experience continued negative outcomes as adults (Moffitt et al., 1996; Odgers et al., 2008), necessitating early intervention. Behavior therapy is one of the most common forms of intervention for CP (Eyberg et al., 2008), yet is ineffective for nearly 30% of youth with CP (Kolko et al., 2009). This may be due to the heterogeneity that is seen in the presentation of CP across and within youth. Affect is a vital factor contributing to this heterogeneity and includes the traits of irritability (IRR) and limited prosocial emotions (LPE). Therefore, the goal of the current proposal is to investigate how these affective factors impact treatment efficacy in youth with CP.

Frick & Morris' (2004) etiological pathways model to CP posits two possible developmental origins, (1) emotion dysregulation or (2) reduced conscience development. The first pathway of emotion dysregulation is characterized by the ability to initiate, maintain, and adjust the intensity and expression of various emotions (for review, see Frick & Morris, 2004). Research has shown that an inability to successfully regulate certain emotions is associated with increased risk for certain forms of psychopathology. For instance, dysregulated sadness is highly related to the onset of internalizing problems. In contrast, disinhibited anger is associated with aggressive and externalizing behavior problems (Eisenberg et al., 2001), which include CP. Indeed, IRR maps onto this first developmental pathway of emotion dysregulation and has been theorized to be the primary affective pathway to CP; children who exhibit aggressive and disruptive behaviors do so due to an inability to properly modulate their anger and frustration (Frick, 2009; Frick & Morris, 2004). IRR is described as overreactions to a situation, such as perceived slights and disagreements, with anger (Wakschlag et al., 2015). Youth who present with elevated rates of IRR show substantial deficits in their executive functioning abilities, particularly inhibitory control (Ellis et al., 2009; Thomson & Centifanti, 2018). These deficits highlight how CP are manifested through two mechanisms underlying IRR. The first is hostile attribution bias, where a child interprets their environment, including neutral stimuli, as threatening or dangerous (Dodge & Coie, 1987). This leads to the second key factor which is reactive aggression, where youth lash out in response to environmental stimuli with impulsive and aggressive behaviors (Brotman et al., 2017). IRR is associated with impairments across various life domains including interpersonal functioning (e.g., parent- and peer-relations) and increased risk for internalizing psychopathology (for review, see Leibenluft & Stoddard, 2013).

In contrast to an inability to regulate emotions, the second avenue to CP revolves around reduced conscience development. More specifically, this pathway suggests that aggressive and antisocial CP are due to deficits in moral reasoning and general empathic concern for the well-being of those around them (Blair, 1999). This lack of empathy is compounded by deficits in guilt and remorse for one's wrongdoings, where children experience diminished negative arousal (e.g., "deviation anxiety;" Frick & Morris, 2004) for their misbehavior, placing them at high-risk to engage in continued antisocial behavior. This combination of reduced empathy, guilt, and remorse is captured by the affective trait of LPE (Frick et al., 2014b), which maps onto the reduced conscience pathway. Often referred to as "callous-unemotional traits" (Frick et al., 2014b), LPE was included in the DSM-5 as a specifier for conduct disorder (American Psychiatric Association, 2013). The inclusion of this specifier in the DSM-5 was based upon nearly two decades of research showing that LPE contributes to highly aggressive and stable CP, including delinquent behavior and later criminal activity (for review, see Frick et al., 2014b). LPE is predominately related to proactive or instrumental forms of aggression (Pardini et al., 2003; Waschbusch & Willoughby, 2008). Further, the presence of LPE has been found to be a precursor to adulthood psychopathy (Christian et al., 1997). In contrast to IRR, only a small subgroup of youth with CP present with LPE (Frick, 2009), and the presence of LPE is inversely associated with internalizing disorders (Frick et al., 1999; Herpers et al., 2014).

The developmental pathway model put forth by Frick & Morris (2004) suggests that the route to CP occurs through either one of these affective traits characterized by emotion dysregulation (IRR) or reduced conscience development (LPE). This highlights the equifinality nature of aggressive and antisocial behavior where various etiological pathways (e.g., IRR, LPE) contribute to the onset of a given outcome, such as CP (Cicchetti & Rogosch, 1996). These traits represent opposite ends of the affective spectrum (over-arousal vs. under-arousal). However, there may be instances where these traits occur within a child as part of situational or state-like dysregulation. For instance, youth with LPE may exhibit more irritability and reactive forms of aggression when their goal-directed activity or aggression is interrupted. Similarly, "...a child who is irritable and thus frequently lashes out at others...may or may not feel guilty after displaying this misbehavior" (Waschbusch et al., 2020, p. 2). The combination of IRR and LPE is associated with elevated impairment, aggression, and increased risk for psychopathology (Byrd, Hawes, Loeber, et al., 2018; Craig & Moretti, 2018; Waschbusch et al., 2020). However, to date, only three studies have investigated this notion which remains exploratory at best.

The development pathways reflected through the affective traits of IRR and LPE indeed contribute to CP in unique manners. This counters the notion of a "one size fits all" intervention protocol for youth with CP. For instance, youth with higher rates of IRR are more amenable to treatments targeting impulsive behavior and anger control. In contrast, the low fearlessness and thrill-seeking disposition seen in youth with higher rates of LPE highlights that these youth may benefit from interventions that aim to promote and establish empathy, particularly through reward-based strategies to capitalize on the child's self-interest (for review, see Frick, 2001). These unique mechanisms underlying IRR and LPE suggest differences in treatment profiles due, in part, to differential response styles to certain treatment components. Rewards (e.g., toys, positive reinforcement) have been found to be effective for all youth with CP, regardless of their level of IRR or LPE (Byrd et al., 2014; Byrd, Hawes, Burke, et al., 2018; Kircanski et al., 2018). In contrast, punishing stimuli (e.g., response cost, time-out) yield different outcomes. More specifically, LPE is associated with reduced responsive to punishing stimuli known as "punishment insensitivity" (Blair, 1999; Frick et al., 2014b). On the other hand, IRR is not associated with said punishment insensitivity and has shown to be effectively targeted using disciplining strategies (Eyberg et al., 2008; Kircanski et al., 2018; Waxmonsky et al., 2021). Thus, it may be that IRR would be predictive of a more favorable treatment profile given that this affective trait is more responsive to rewarding and punishing treatment aspects. In contrast, LPE may predict a limited treatment profile, as it is more responsive to rewards, but less responsive to punishing and discipling aspects. Therefore, more effective and personalized interventions are needed, particularly behavior therapy, a highly effective and common treatment modality for CP (Eyberg et al., 2008). Such personalized forms of intervention would be more effective in targeting the distinct treatment mechanisms underlying CP.

While IRR and LPE may differ in how they relate to treatment efficacy, to date, no work has examined these two affective traits simultaneously within the context of a treatment study. In fact, there appears to be virtually no empirical work examining how the specific trait of IRR impacts treatment efficacy for CP. This is a significant limitation considering that IRR is the primary pathway to CP, yet LPE has received more empirical attention despite being observed in only a minority of youth with CP (Frick, 2009). Further research is also needed to parse out unique aspects of treatment efficacy. First, the terms treatment "response" and treatment "outcome" are often conflated despite representing distinct aspects of treatment efficacy. Next, many treatment studies have utilized pre-post designs which inadvertently ignore change *during* treatment. Lastly, there is a dearth of research examining "in the moment" change and reaction to time-out, a specific behavior therapy contingency that is long enough to conduct such investigation. The current study aims to examine how the affective traits of IRR and LPE uniquely predict treatment efficacy in youth with CP, with treatment efficacy defined as (1) treatment "response" and treatment "outcome" (2) trajectories of symptom and impairment change, and (3) reactions to time-out. These aims will be examined within the context of the Summer Treatment Program (STP).

1.1 Summer Treatment Program (STP)

The STP is a manualized, evidenced-based program for youth presenting with impairing hyperactive, disruptive, and aggressive behaviors. The STP is a well-established and highly effective treatment for youth with Attention-Deficit/Hyperactivity Disorder (ADHD) and has also shown to be effective for youth with CP (Pelham Jr. et al., 2010; Pelham Jr. & Hoza, 1996). The STP is a unique treatment protocol as it targets the distinct constructs of symptoms *and* impairment (Pelham Jr. & Hoza, 1996). Symptoms are behavioral manifestations of a disorder (e.g., "often argues with authority figures" [oppositional defiant disorder]) whereas impairment reflects how much of a problem the symptoms of a disorder have upon the individual and others (Fabiano et al., 2006). The

STP employs several behavior modification techniques which are used to (1) reduce impairment that children experience from their disruptive symptoms and (2) promote prosocial behaviors and reduce disruptive and antisocial behaviors. These techniques include a token economy system, daily report cards, positive reinforcement, parent management training, and time-out. This intensive, multiweek, treatment package is held during the summer months and runs from 8 a.m. to 5 p.m., Monday through Friday. This culminates in nearly 400 hours of intensive behavior therapy, which is equivalent to nearly seven years of weekly therapy in an outpatient setting (Pelham Jr. et al., 2017).

The STP is an ideal and well-designed setting to examine the study aims of (1) treatment response and treatment outcome, (2) trajectories of change, and (3) moment-to-moment contingency analysis. First, its repeated-measures design permits for constant monitoring and recording of all child behaviors throughout treatment, allowing for feasible examination of both treatment response and trajectories of change. Next, data regarding child behavior are collected prior to and after treatment, allowing for analysis of treatment outcome. Lastly, the STP uses several contingency systems, such as a standardized time-out protocol, which provides raters the ability to record a child's moment-to-moment behavioral and emotional reactions. Given the STP's focus on symptoms and impairment, it is imperative to examine how IRR and LPE uniquely predict *both* these distinct outcomes. This is particularly salient when considering that much of the prior work has primarily focused on symptoms (e.g., Hawes et al., 2014; Webster-Stratton & Hammond, 1997), and impairment is a necessary element across psychopathology (American Psychiatric Association, 2013).

1.2 Treatment "Response" and Treatment "Outcome"

First, the terms treatment "response" and treatment "outcome" are often used interchangeably yet are distinct concepts. Treatment response describes *the magnitude or amount of change* from pre- to post-treatment while treatment outcome represents the *likelihood of normalization* at post-treatment (e.g., Lindhiem et al., 2012). Using this distinction, prior work has found that youth with CP experience a large treatment response, but poor treatment outcomes (Reyno & McGrath, 2006). A similar pattern was seen in youth with high levels of both CP and LPE (Bansal et al., 2019). Indeed, there may be a ceiling effect where higher pre-treatment severity of CP allows for more change to occur, warranting the need to control for baseline CP. Further, the theoretical underpinnings of IRR and LPE may identify how treatment response and outcome may differ as a function of affect. For instance, youth with CP and IRR may experience a greater treatment response and better treatment outcomes relative to youth with CP and LPE. This is reinforced by the notion that many behavior therapy contingencies are designed to broadly target emotion dysregulation, particularly impulsive and anger-driven reactions to situations, consistent with the IRR presentation of CP (Eyberg et al., 2008; Frick, 2001).

1.3 Trajectories of Symptom and Impairment Change

Next, much of the extant treatment literature has utilized pre-post treatment designs to assess the amount of change that occurred, which implicitly assumes that change is a linear process, regardless of the presenting concern or the population. The use of multiple assessment points throughout treatment would highlight *when* interventions engender change and could inform adjustments to the protocol to yield even greater change (e.g., duration). Little work has examined how trajectories of change apply to CP, with only two studies highlighting a curvilinear decline in CP over time (Lindhiem & Kolko, 2010, 2011). Further work is needed to determine how IRR and LPE impact the timing of symptom and impairment change. Such work could greatly enhance clinical practice and treatment design, such as personalizing the duration of the intervention to better fit the child's affective profile. Given that youth with low levels of LPE are most amenable to behavior therapy, it may be that IRR would predict a greater slope in change relative to LPE.

1.4 Time-Out

Lastly, further work is needed to isolate the nuanced effects of distinct modification techniques apart from the larger treatment protocol (Chorpita & Daleiden, 2009). Time-out is an ideal contingency to examine these nuanced effects, particularly because it is the only component that is long enough to allow for moment-to-moment analysis of emotional and behavioral reactions. This would help further evaluate how effective time-out is for youth with CP and different affective profiles. Time-out targets problematic behaviors by removing the child from a reinforcing situation following an inappropriate behavior and placing them in a non-reinforcing environment (Fabiano et al., 2004). Consistent with the "punishment insensitivity" paradigm, time-out is generally less effective for youth with higher LPE, as these youth are less emotionally reactive to time-out (Bansal, Haas, et al., 2020; Dadds & Rhodes, 2008; Hawes & Dadds, 2005) but also exhibit greater rates of disruptive behaviors during time-out (e.g., aggression, cursing; Bansal, Haas, et al., 2020; Haas et al., 2011). In one study, youth with CP and high LPE were most likely to exhibit negative behaviors when emotions were rated as "amused" relative to any other emotion (e.g., sad, angry; Bansal, Haas, et al., 2020). It may be that LPE would be associated with higher rates of negative behaviors that are exhibited out of enjoyment or amusement. In contrast, IRR may be associated with greater responsiveness to time-out; IRR may predict higher initial rates of negative behaviors exhibited due to anger but would become more amenable to the time-out process over time.

1.5 The Current Study

IRR and LPE have been highly influential in elucidating the heterogeneity of CP. However, little work has examined how these unique affective traits predict symptoms of CP and overall impairment across the distinct aspects of treatment efficacy. In the context of an intensive STP, the current study aims to examine how (1) treatment response and treatment outcome, (2) trajectories of change, and (3) reactions to time-out all differ as a function of IRR and LPE. After controlling for baseline CP severity in aims 1 and 2, it was hypothesized that: (1) IRR would predict a large treatment response and positive treatment outcomes, while LPE would predict a small treatment response and poor treatment outcomes; (2) IRR would predict a curvilinear reduction in symptom/impairment change, while LPE would predict a flat pattern of symptom/impairment change; and (3) IRR would predict negative behaviors associated with anger or sadness during time-out, and that these behaviors would decrease over time, while LPE would predict negative behaviors associated with amusement, and that these behaviors would not decrease over time.

CHAPTER 2. METHODS

2.1 Participants

Participants were enrolled in an STP held in a northeastern state of the United States in 2018 and 2019. The sample included 49 children (grade: *M* = 3.22, *SD* = 1.49), comprised of 35 males (71.4%) and 14 females (28.6%) ranging from seven to 12 years of age (*M* = 9.02, *SD* = 1.37). Thirty-eight children (77.6%) identified as Caucasian, five children (10.2%) as African American, and six children (12.2%) as other (Asian, Native American, Biracial). The average full-scale IQ, as measured by the Weschler Intelligence Scale for Children-5th edition (Wechsler, 2014) was estimated to be 96.09 (*SD* = 14.80). Most female caregivers were biological mothers (*n* = 33, 67.3%), while other female caregivers included adoptive mothers (*n* = 8, 16.3%), grandmothers (*n* = 4, 8.2%), stepmothers (*n* = 2, 4.1%), fosters mothers (*n* = 1, 2.0%), and other (*n* = 1, 2.0%). Most male caregivers (*n* = 34) were biological fathers (*n* = 23, 67.6%), with the remaining caregivers identifying as adoptive fathers (*n* = 7, 20.6%), stepfathers (*n* = 2, 5.9%), foster fathers (*n* = 1, 2.9%), and grandfathers (*n* = 1, 2.9%). The average income as reported by female caregivers (*n* = 49) ranged from approximately $50k to $59k while the average income as reported by male caregivers (*n* = 33) ranged from approximately $60k to $69k.

Exclusion criteria included (1) age younger than 7 or age older than 12.99 on the first day of treatment; (2) full scale IQ < 70; (3) past or current diagnosis of pervasive developmental disorders, active psychotic disorders, sexual disorders, or eating disorders; (4) any physical or medical condition that would preclude participation in the STP; and (5) use of medication other than stimulants. Of the 49 youth, 38 (77.6%) met criteria for CP, including 16 youth (32.7% of the full sample) who met diagnostic criteria for *both* oppositional defiant disorder and conduct disorder. All youth with CP (*n* = 38) also met criteria for Attention-Deficit/Hyperactivity Disorder (ADHD), primarily ADHD combined-type (*n* = 35, 92.1%). Eight youth (16.3%) met diagnostic criteria for ADHD without CP, while three youth (6.1%) did not meet diagnostic criteria for any disruptive behavior disorder and were considered controls. Thirty-eight youth (77.6%) were on stimulant medication during treatment.

2.2 Procedures

2.2.1 Recruitment

This study was approved by a university-based Institutional Review Board. Written and informed consent was obtained from all caregivers and verbal assent was obtained from children. Similar to prior STPs (e.g., Bansal et al., 2019), study participants and families were recruited through a variety of methods including fliers sent to schools and mental health professionals, and posted in public locations (e.g., libraries, community centers, grocery stores), and advertisements on TV and the radio. Fliers were targeted towards youth who displayed attention and behavior problems. primarily operationalized as ADHD. Caregivers who were interested in the study contacted the lab and were given a brief phone screening evaluation. After this, if the interested caregiver was eligible and still interested, lab coordinators sent out caregiver and teacher ratings to be completed and returned; this battery of questionnaires served as the pre-treatment ratings. The battery of ratings given at pre-treatment differed from the post-treatment ratings. Given this, the magnitude of change in CP and impairment from *pre*-treatment to *post*-treatment could not be explicitly examined. Instead, treatment response analyses used data obtained *during* treatment. Thus, from here on forward, treatment response was operationally defined as the magnitude of change from the *beginning* of treatment (i.e., week 1 of the STP) to the *end* of treatment (i.e., week 5 of the STP). Diagnoses were determined using DSM-5 criteria. Following best practice guidelines for ADHD and CP (McMahon & Frick, 2010; Pelham Jr. et al., 2005), diagnoses were determined based upon symptom and impairment levels as evaluated using multiple sources of information including caregiver ratings, teacher ratings, and a structured clinical interview administered to caregivers. Final diagnoses were made by doctoral-level clinicians.

In total, 73 children were recruited between the 2018 and 2019 STPs. Of these 73 children, seven did not attend the treatment program. Of the remaining 66, 11 children were "repeat" participants, as they participated in both the 2018 and 2019 STPs, leading to 22 total data entries for these 11 youths. For these repeat participants, only their 2018 data was kept to ensure that all children were analyzed during their first exposure to the treatment protocol. Removal of the 2019 data for these repeat participants led to a sample of 55 children. Lastly, six youth were removed from analysis because they were younger than seven years old. The STP protocol has been adapted for younger children, hence why they were not included in analysis (i.e., STP-PreK; Graziano et al., 2014). In total, this led to the final sample of 49 children who attended a single year of treatment and were the appropriate age; these children were included in analyses. Pre-treatment parent ratings were available for all participants (n = 49) while pre-treatment teacher ratings were available for 21 participants (42.9%). No differences were seen between youth whose data were included in analyses vs. those whose data were not included in analyses on several factors including age (p = .054), grade (p = .12), full scale IQ (p = .99), IRR total score (p = .23), LPE total score (p = .051), pre-treatment CP count score (p = .18), and pre-treatment impairment (p = .94).

2.2.2 Treatment

The STP was conducted for five weeks, Monday through Thursday, 8 a.m. to 4 p.m., and 8 a.m. to 1 p.m. on Fridays. Several behavioral modification techniques were employed throughout the STP including a comprehensive token economy system (e.g., “point system”), a “point store” where children could exchange their accumulated points throughout the week for a toy, individualized daily report cards, constant use of positive reinforcement, social skills and problem-solving discussion groups, and time-out (from positive reinforcement). A doctoral level clinician and advanced graduate student, both of whom had significant experience with the STP, extensively trained the treatment staff to ensure that the manualized treatment protocol was accurately and effectively implemented. Participants were separated into groups consisting of 12-15 similarly aged peers. Each group was monitored by five undergraduate counselors and a graduate-level lead counselor.

Mondays through Thursdays, children participated in three hours of structured sports skills (e.g., soccer, kickball, softball, and swimming) and two hours of academic learning activities (e.g., math, reading). Additional activities included group social skills and problem-solving skills, which occurred at the beginning of each new activity, along with lunch and recess. On Fridays, children engaged in group discussions and were then separated into one of three levels for the remaining two hours of day: Level 1 or “Fun

Friday" which consisted of rewarding activities (e.g., pizza and/or pool party) and a relaxed treatment protocol; Level 2 (e.g., "regular camp group") where children engaged in a typical and structured day of treatment; and Level 3 (e.g., "chores group") where children engaged in punishing activities, such as doing chores or writing sentences. Children were placed in these groups based upon established criteria rating each child's behavior throughout the week. During the 2018 STP a weekly, hour-long parent training skills group, led by a doctoral level clinician, was provided at no cost to the participants' caregivers. Across the five treatment weeks, the average number of attended parent training groups was 2.95 (*SD* = .99). Given the sporadic attendance of this group during the 2018 STP, the parent training skills groups was not offered during the 2019 STP. Caregivers and counselors completed weekly rating scales regarding child impairment and improvement. Post-treatment ratings were also completed by caregivers and counselors but rather than repeating the same measures at baseline (i.e., pre-treatment), the post-treatment ratings assessed child improvement and satisfaction with the STP.

2.2.3 Time-Out

In addition to examining the effects of the general treatment protocol, the current study explored the effects of time-out. Time-outs in the current study followed the typical STP format (Fabiano et al., 2004). Briefly, children received a time-out after exhibiting one of the following behaviors, repeated noncompliance, intentional aggression (towards peer or staff), and intentional destruction of property. As seen in Fabiano et al., (2004), time-out duration was contingent upon the child's behavior (see Figure 1). If the child served the first half of their time-out appropriately (e.g., first five minutes of the assigned ten-minute time-out), their time-out would end. If the child exhibited any negative behaviors during the time-out process, their time-out length was escalated (e.g., 10-minute time-out escalated to a 20-minute time-out). Treatment counselors monitored and recorded all behaviors on the corresponding time-out log. Specific to the current study, at the end of each time-out children and counselors completed a variety of items assessing the child's mood during time-out, and possible functions of the child's negative behavior. These post-time-out items were developed and implemented just prior to the start of the 2019 STP. Therefore, analyses only focused on the time-outs that occurred during the 2019 STP.

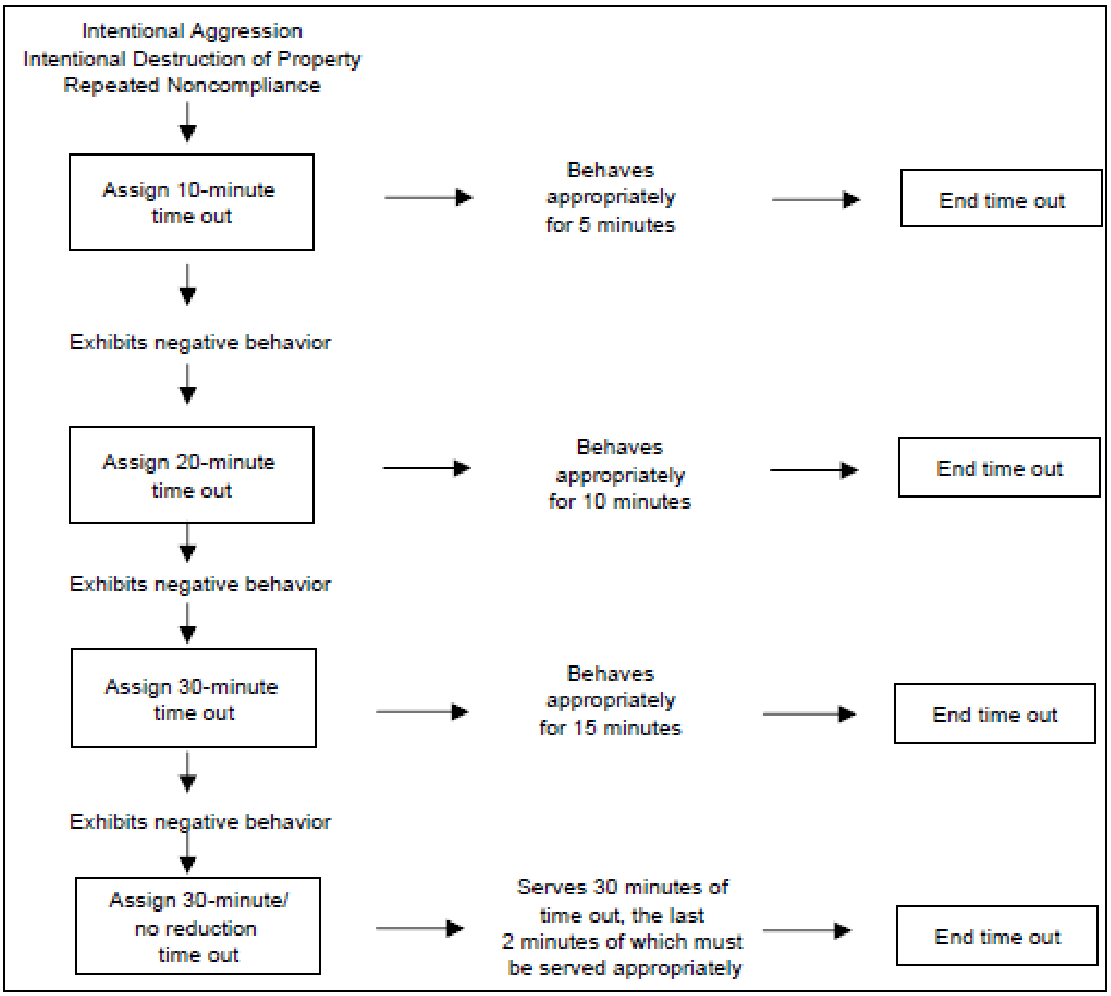

Figure 1. Sequence of Establishing Time-Out Duration in the STP.

2.3 Measures

2.3.1 IRR

IRR was assessed at pre-treatment using the parent-report version of the Affective Reactivity Index (ARI; Stringaris et al., 2012). The ARI consists of seven items rated on a three-point, 0 (*Not true*) to 2 (*Certainly true*) scale. The first six items assess IRR and add up to a total score, while the seventh item assesses impairment from IRR and is not counted in the total score. Items on the parent-report version, as rated by female caregivers (α = .92) and male caregivers (α = .92), showed excellent internal consistency. Parent-report scores were determined by taking the highest score reported between female and male caregivers, a method that has empirical support in prior work (Stringaris et al., 2012). Pre-treatment parent-report scores showed an average IRR score of 5.65 (*SD* = 3.56). Scores ≥ 3 on the parent-report ARI signify clinically elevated rates of IRR (Kircanski et al., 2017). Using this approach, 35 children (72.9%) were classified as having clinically significant levels of IRR. Total scores on the IRR were used as the predictor variable for the statistical models. The ARI is a valid and reliable measure of IRR during childhood (Stringaris et al., 2012).

2.3.2 LPE

LPE was assessed at pre-treatment using the parent-version of the ICU (Frick, 2004). The 24 items are rated on a 0 (*Not at all true*) to 3 (*Definitely true*) scale, with 12 negatively worded items ("Shows no remorse when he/she has done something wrong") and 12 positively worded items ("Works hard on everything"). Consistent with past research (Bansal, Babinski, et al., 2020), the 12 positively worded items were reverse scored and then summed with the 12 negatively worded items to create a total score. The LPE total score was used as a predictor for analyses based upon its wide use and a recent meta-analysis supporting use of a total score (Ray & Frick, 2018). Items on the parent-report version of the ICU, as rated by female caregivers (α = .88) and male caregivers (α = .86), showed good internal consistency. Parent-report scores were determined by taking the highest score reported between female and male caregivers, a method that has been used in prior studies of LPE (Frick et al., 2003). Pre-treatment parent-report scores demonstrated an average total LPE score of 31.59 (*SD* = 9.84). Based upon established norms using a representative sample of U.S. children (Bansal, Babinski, et al., 2020), this average LPE score is equivalent to a *t*-score of 56; approximately 16% of the sample (*n* = 8) had an LPE *t*-score at or above 65 (LPE total score ≥ 42).

2.3.3 Disruptive Behavior Disorders

Parents completed the Disruptive Behavior Disorder Rating Scale (DBDRS; Pelham Jr. et al., 1992) prior to the start of treatment. The DBDRS assesses for ADHD (18 items), oppositional defiant disorder (eight items), and conduct disorder (15 items). Items are scored on a 0 (*Not at all*) to 3 (*Very much*) scale. Pre-treatment ADHD and CP scores were computed and included in analyses as covariates. The 18 items assessing ADHD – nine for inattentive symptoms (α = .85) and nine for hyperactive/impulsive symptoms (α = .87) – were summed together to create a total ADHD score (α = .91). The six oppositional defiant items (α = .91) and 15 conduct disorder items (α = .64) were summed together to create a total CP score ($\alpha = .89$). The average pre-treatment ADHD score was 1.95 ($SD = .69$) and the average pre-treatment CP score was .59 ($SD = .41$).

2.3.4 CP

Observational data obtained from the counselors' point system records was used to compute daily frequencies of CP, which was used as the first outcome variable. This observational data was recorded daily throughout the five weeks of the STP. A CP sum score was computed by summing together the behavioral categories of intentional aggression, intentional destruction (of property), lying, and stealing (Pelham Jr. et al., 2000). First, a daily CP score was created by summing these categories. Then, the daily CP scores were summed to create a weekly CP sum score. Given that this score was based upon behavioral observations, this score was not normed. This weekly sum-score only included data from Monday to Thursday; this was done because the treatment protocol was not uniformly administered to all children on Fridays, as some children experienced a relaxed protocol whereas others had a typical camp day. This yielded five CP weekly scores for the five weeks of treatment. Rather than using a weekly average score, analyses used a sum score because CP can be considered a non-negative count variable as seen in prior research (Monuteaux et al., 2006). Prior STP studies have employed observational data as outcomes (Miller et al., 2014; Pelham Jr. et al., 2000; Waschbusch et al., 2007), and these data have been found to be reliable measures of CP (e.g., Pelham Jr. et al., 2000). Descriptive statistics for the CP sum count score for weeks 1 through 5 can be found in Table 1.

2.3.5 Impairment

Parents completed the Impairment Rating Scale (Fabiano et al., 2006), which served as the second outcome variable, prior to the start of treatment and during each of the five weeks of the STP. This measure consists of seven items assessing impairment across various functional domains: peers, siblings, parents, family, self-esteem, overall adjustment, and need for treatment. These items are rated on a visual-analogue scale ranging from 0 (*No problem*; *No need for more treatment services*) to 6 (*Extreme problem*; *Extreme need for more treatment services*). These items were averaged together to create an overall impairment score; the pre-treatment score was used as a baseline measure of impairment and included in analyses as a covariate, while five weekly impairment scores were created to reflect the five treatment weeks. Items on the parent-report version of the IRS, as rated by female caregivers (α = .87) and male caregivers (α = .90), showed good to excellent internal consistency. Parent-report scores were determined by taking the highest score reported between female and male caregivers. The IRS has been found to be a valid and reliable measure of impairment in school-aged youth with disruptive behavior disorders (Fabiano et al., 2006). Table 1 presents the descriptive statistics for the IRS across the five weeks of treatment.

2.3.6 Time-Out

Counselors completed the investigator-designed time-out log during the time-out process (see Appendix A). This log captures various time-out characteristics such as time-out reason, duration of physical management, and a litany of negative behavior categories (e.g., aggression, cursing, complaining). At the end of each time-out, children and counselors completed items assessing the child's level of (1) anger, (2) amusement, and (3) guilt during time-out rated on a 1 (*Not at all*) to 5 (*Extremely*) scale. Counselors also completed four items regarding possible motivation for the child's behavior during time-out (e.g., the child was trying to make their counselors and/or peers mad and/or laugh) which were rated on a 1 (*Strongly disagree*) to 5 (*Strongly agree*) scale. As aforementioned, these items were only developed for the 2019 STP.

Several outcome variables were created for analyses. First, continuous outcome scores included time-out duration, physical management duration, and total number of negative behaviors during time-out. Next, categorical outcome variables included the child-rated and counselor-rated anger, amusement, and guilt items along with the counselor-rated motivation items. Child- and counselor-rated items were computed separately. All scores were created by taking the average score of each variable across the five treatment weeks (e.g., average time-out duration was calculated for all five weeks). This led to five mean scores for all continuous and categorical variables. Unlike the point system data for CP where scores were averaged from Monday through Thursday, the time-out procedure was consistently implemented throughout treatment. Therefore, average scores for time-outs were computed Monday through Friday.

Table 1. Descriptive Statistics for CP and IMP Outcomes.

Week	Statistic	CP	IMP
1	Mean (SD)	.80 (2.37)	2.33 (1.24)
	Median	.00	2.33
	Skew	3.29	-.01
2	Mean (SD)	1.35 (5.93)	2.57 (1.21)
	Median	.00	2.80
	Skew	6.52	-.30
3	Mean (SD)	4.00 (15.16)	2.28 (1.22)
	Median	.00	2.42
	Skew	4.75	-.04
4	Mean (SD)	9.31 (49.10)	2.28 (1.27)
	Median	.00	2.33
	Skew	6.47	.15
5	Mean (SD)	4.20 (15.34)	2.26 (1.23)
	Median	.00	2.08
	Skew	4.18	.29

Note. CP = Conduct Problems as measured by behavioral observations; IMP = Impairment as measured by the Impairment Rating Scale (Fabiano et al., 2006); CP measured using sum counts, IMP measured using average scores.

2.4 Analytic Plan

Prior to formal analysis, several data checks were conducted including missing data (prevalence and type), normality of distributions, possible outliers, and relations amongst study variables. Following this, data were examined to determine if any youth dropped out of treatment; independent samples *t*-tests were conducted to determine if these youth differed from youth who completed treatment. These preliminary analyses were conducted in SPSS v27. All primary analyses were all conducted in SAS v9.4. All missing data was first handled using estimation procedures contingent upon the SAS procedure (e.g., PROC MIXED, PROC GLIMMIX). In all models, IRR and LPE were included as continuous measures rather than categorical variables for several reasons. First, the distribution of both IRR and LPE in the current sample appeared to approach a normal distribution. Next, these constructs have been found to be continuously distributed in clinical and community samples (Hudziak et al., 2007). Lastly, continuous predictors have been found to be more statistically powerful than categorical predictors (Altman & Royston, 2006).

Models also examined several theoretically relevant factors as covariates. First, models controlled for pre-treatment CP and IRS scores to ensure that change in these outcome variables was not due to a ceiling effect (e.g., higher scores at pre-treatment contributing to greater change). Next, a sum score parent-training variable was created reflecting the total number of parent training groups that were attended by each child's caregiver. This variable was included because caregivers in the 2018 STP received one-hour of parent training skills while caregivers in the 2019 STP did not (all caregivers in 2019 received a '0' on this variable). Models also controlled for ADHD (continuous scores) as this has been shown to be a longitudinal predictor of CP (Pardini & Fite, 2010). Lastly, two dichotomous covariates were included: (1) medication status (yes vs. no) because this differed across children and (2) sex (male vs. female) because CP has shown to differ between males and females (for review, see Waschbusch, 2002). All continuous predictors and covariates were mean centered prior to analysis.

2.4.1 Aim 1 – Treatment Response

Aim 1 examined treatment response, which reflected the magnitude of change in CP and impairment from the beginning of treatment (i.e., week 1) to the end of treatment (i.e., week 5). For the treatment response analyses, two separate models were conducted for the two dependent variables, both of which were measured at week 5, (1) CP behavioral observation ratings and (2) impairment ratings on the parent-reported IRS. Assumptions of normality were violated by the CP outcome which showed a positively skewed distribution consistent with overdispersion (i.e., sample variance > sample mean; Gardner et al., 1995). Negative binomial regressions through PROC GENMOD with a log link function accommodated these skewed, non-negative count data. As shown in prior work (Huang & Cornell, 2012; Jaffee et al., 2013), the coefficients obtained from these models are interpreted as the difference in the log of expected CP for every one-unit change in the predictor variable. In contrast, the IRS outcome showed a relatively normal distribution and was analyzed using linear regression through the PROC REG function in SAS. For both outcomes, models were constructed in a stepwise fashion: Model A, where the outcome was regressed onto all theoretically relevant covariates and then Model B, where the outcome was regressed onto all covariates and added the effects of IRR, LPE, and their interaction (IRR*LPE). It was hypothesized that IRR would predict a large treatment response while LPE would predict a small treatment response.

Importantly, because the sample only consisted of three (6.1%) youth who were considered typically developing controls (i.e., no diagnosis), treatment *outcome* was not able to be reliably analyzed and was dropped from the study's aims.

2.4.2 Aim 2 – Trajectories of Change

Aim 2 explored how IRR and LPE predicted trajectories of CP and impairment across the five treatment weeks. CP was analyzed using multilevel modeling through PROC GLIMMIX using a negative binomial distribution and log link with random effects and a variance components matrix structure. Laplacian methods were used to estimate models due to (1) its efficiency and flexibility with a wide range of models (e.g., non-nested models) and (2) limitations with default pseudo-likelihood estimation methods (e.g., biased results for small samples, absence of a true log likelihood; Schabenberger, 2007). The IRS outcome was analyzed using multilevel modeling through PROC MIXED with random intercepts and an unstructured covariance matrix. Kenward-Roger corrections to degrees of freedom were applied to all MIXED models to account for the sample size (McNeish, 2017).

Model taxonomy for both PROC GLIMMIX and PROC MIXED was as followed. First, three unconditional models were constructed. Model A (i.e., means model), allowed for random intercepts; Model B (i.e., growth model) included random intercepts along with the effect of time, which was represented by the week of treatment and was centered at 0 indicating the beginning of treatment. Next, Model C included a quadratic term (week*week) to account for possible curvilinear trajectories of CP as shown in prior work (Lindhiem & Kolko, 2010). The first conditional model was Model D, where the effects of all covariates were included on the intercept; covariate effects were not included at the slope given the number of parameters in relation to the overall sample size. Finally, Model E included all covariates and effects of IRR, LPE, and their interaction, along with separate predictor*week interactions (e.g., IRR*week).

For nested models, χ^2 difference tests were conducted to determine whether full models showed significant improvement over restricted models. In instances where models were not nested, model comparison was done using Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), where models with smaller AIC and BIC values signified better fit (Burnham & Anderson, 2004). IRR was expected to predict a curvilinear reduction in CP and impairment change, while LPE would predict a flat pattern of CP and impairment change.

2.4.3 Aim 3 – Time-Out

Aim 3 investigated how IRR and LPE predicted changes in behavioral and emotional reactions to time-out. Of the 17 youth who received a time-out in 2019, several were dropped from analyses. First, time-outs from three youth had no identifying information and could not be linked back to any baseline data. Next, time-outs from four youth were excluded as these youth had participated in both years of the STP. Lastly, time-outs from two children were excluded, one of whom was younger than seven years old during treatment, and the other who was missing data on all post-time-out items. This led to a final sample of eight youth with complete time-out data. Given this sample size, and the unbalanced nature of the time-out dataset, only between-persons analyses were conducted.

First, basic descriptive statistics of each time-out (e.g., duration, reason, etc.) were presented for each youth. Following this, PROC GENMOD, with a negative binomial distribution and log link function, was used to analyze the negative behavior count variable given its skewed, non-negative nature. Next, the 10 items assessing the child's emotional state during the time-out process were analyzed: three child-report items assessing affect (i.e., anger, amusement, and guilt); three counselor-reported items assessing affect (i.e., anger, amusement, and guilt); and four counselor-reported items assessing a child's motivation (i.e., making the counselors mad, making the counselors laugh, making their peers mad, making their peers laugh). Each of these items were averaged across the duration of camp into a single variable, yielding 10 average scores. Several of these scores presented with elevated skewness (≥ 1): child-reported anger, child-reported amusement, child-reported guilt, counselor-reported anger, counselor-reported guilt, counselor-reported motivation (i.e., making counselors mad). The skew of these variables was handled using log10 transformations to improve normalcy of these outcomes. These updated variables were included in the PROC REG analyses. For each outcome, Model 1 consisted of covariates only (i.e., medication status and pre-treatment ratings of CP, impairment, and ADHD); parent training and sex were not included, given that no parent training groups were offered during 2019, and every youth who was included within these time-out analyses was a male participant. Model 2 consisted of the covariates, IRR, and LPE. Due to the small sample size ($n = 8$) and given considerations in testing moderation (Fairchild & MacKinnon, 2009), interactions were not included in these regression models.

CHAPTER 3. RESULTS

3.1 Preliminary Analyses

Examination of attendance records showed that all 49 youth completed all five weeks of treatment. Thus, no youth who were included in analysis dropped out of treatment prematurely. Of the 25 treatment days, youth attended an average of 20.27 (81.1%) camp days (SD = 2.62) across the 2018 and 2019 STPs; common reasons for missing camp included vacation or child illness. Next, amount and type of missing data was examined across the study variables. All youth had complete data on all predictors and covariates, as well as on the CP outcome variable across all five weeks. Minimal data loss was seen for the IRS across the five weeks: week 1 = 8.1%, week 2 = 0%, week 3 = 2%, week 4 = 22%, and week 5 = 10%. Given the higher rates of missing data for the week 4 IRS score, analysis was conducted to determine type of missing data. A dummy variable was created where 0 = no missing week 4 data and 1 = missing week 4 data. This variable was correlated with several demographic (e.g., age, sex grade) and predictor variables (e.g., IRR, LPE, CP scores). No significant correlations were observed (*p*-value ranges .10 - .73), suggesting no systematic patterns in the missing IRS data at week 4.

3.2 Aim 1 – Treatment Response

3.2.1 CP

Examination of treatment response for the week 5 CP behavioral observation outcome variable was done using negative binomial regressions in PROC GENMOD. Fixed effect estimates are presented in-text, while complete results, including exponentiated coefficients, are presented in Table 2. Model A, the covariates-only model, found a significant effect of sex (*b* = 3.36, $p < .05$), suggesting that males presented with significantly higher levels of CP at week 5. Next, Model B, the covariates and predictors model, observed significant effects pre-treatment CP (*b* = 8.07, $p < .01$), suggesting that higher rates of baseline CP predicted higher rates of week 5 CP. Further, there was a significant main effect of IRR (*b* = -.92, $p < .01$), suggesting that for every one-unit increase in pre-treatment IRR, the log count of CP was expected to decrease by .92 units. In other words, pre-treatment IRR was predictive of week 5 CP, with higher levels of IRR predicting reductions in CP at week 5. Neither LPE (*b* = -.07, $p = .24$) nor the IRR*LPE interaction (*b* = -.01, $p = .47$) were significant.

Table 2. Fixed Effects Estimates and Exponentiated Coefficients for CP Treatment Response.

Model	Variable	*b*	Exp(*b*)
1	Intercept	-2.66	.07
	Pre-treatment CP	.88	2.41
	Sex (Male)	3.36*	28.71
	Parent-Training	.22	1.24
	ADHD	2.19	8.90
2	Intercept	-4.77*	.01
	Pre-treatment CP	8.07**	3187.21
	Sex (Male)	1.80	6.04
	Parent-Training	.48	1.62
	ADHD	2.71	15.08
	IRR	-.92**	.40
	LPE	-.07	.93
	IRR*LPE	.01	1.01

Note. CP = Conduct Problems; ADHD = Attention-Deficit/Hyperactivity Disorder; IRR = Irritability; LPE = Limited Prosocial Emotions; *$p < .05$; **$p < .01$.

3.2.2 Impairment

Examination of treatment response for the week 5 impairment outcome variable on the parent-reported IRS was done using linear regressions in PROC REG. Ultimately, neither Model A (covariates-only: $F = 1.47$, $p = .22$) nor Model B (covariates and predictors: $F = 1.07$, $p = .41$) were significant, suggesting that the covariates nor predictors of interest exhibited any significant effect upon week 5 impairment.

3.3 Aim 2 – Trajectories of Change

3.3.1 CP

Examination of CP trajectories across the five weeks of treatment was done using PROC GLIMMIX. Models of CP trajectories using a variance components matrix structure did not produce a covariance estimate between the intercept and slope, nor did they provide level-1 residual estimates. To obtain these estimates, additional models were fit using certain methods (i.e., unstructured matrix structures, Cholesky algorithms, specifying parameter estimates, non-Laplacian estimation methods). These modifications led to model convergence issues and not positive definite matrix structures. Thus, models used variance components structures which produced variance estimates for the intercept and slope separately but were unable to produce the covariance estimate between these parameters and the level-1 residual estimate.

Complete results, including exponentiated coefficients and χ^2 difference tests, are presented in Table 3 with fixed effect estimates presented in-text. There was a significant effect of the intercept ($b = -3.13$, $p < .01$) in Model A (i.e., unconditional means model), suggesting that week 1 CP scores were significantly different from 0. The intercept in Model B (i.e., unconditional growth model) was also significant (b = -4.04, p < .001), once again suggesting that week 1 CP scores were significantly different from 0. The variance parameter estimate for the intercept was significant for both models (Model A estimate = 11.85, p = .01; Model B estimate = 11.89, p = .01) indicating that there were significant differences in the starting point for CP across youth. Further, the slope was significant in Model B (b = .38, p < .01), signaling that for each week of treatment, the log count of CP was expected to increase by .38 units. In other words, CP appeared to *increase* over the course of treatment. The variance parameter estimate for the slope in this model was not significant (estimate = .001, p = .50) suggesting no differences in how CP changed over time across youth. Figure 2 shows the mean score of CP across the five treatment weeks.

Model C included a quadratic effect (week*week); while significant effects were seen for both the intercept and slope in this model, the quadratic effect was not significant (b = -.14, p = .10). χ^2 difference tests found that the addition of the quadratic effect did not lead to a better fitting model (p = .11) and was excluded from all subsequent models. Model D included random effects of the intercept and slope along with all theoretically relevant covariates. The intercept of this model remained significant; a significant effect of sex upon the intercept was also observed (b = 2.82, p = .04), suggesting that males showed higher levels of CP at week 1 relative to females. The effect of treatment week remained significant (b = .36, p < .01), indicating that CP *increased* over the course of treatment. More specifically, the log count of CP was expected to increase by .36 units for every week of treatment. Model improvement was determined using AIC/BIC indices since Models C and D were not nested. For Model C, AIC = 493.3 and BIC = 505.2 while for Model D, AIC = 482.5 and BIC = 501.2, suggesting that Model D provided a better fitting model.

Lastly, Model E included all covariates along with the predictors of interest (IRR, LPE). The model G matrix was not positive definite. This issue was remedied by removing the random effect of time because the estimated effect of this parameter was negative. The model subsequently converged and significant effects were observed for the intercept, sex, and slope (Table 3). However, neither IRR nor LPE exerted any significant effect upon the intercept nor slope. Further, χ^2 difference tests found that this model was not a significant improvement over Model D ($p = .69$), suggesting that Model D was the best fitting and most parsimonious model.

Table 3. PROC GLIMMIX Effects of CP Trajectories.

	Model A	Model B	Model C	Model D	Model E
Initial Status					
Intercept	-3.13 (.04)**	-4.04 (.02)**	-4.34 (.01)***	-5.08 (.01)**	-5.43 (.004)**
Pre-tx CP	--	--	--	.48 (1.61)	2.10 (8.17)
Sex (Male)	--	--	--	2.82 (16.83)*	2.60 (13.43)*
Medication	--	--	--	-1.12 (.33)	-.24 (.78)
Parent Training	--	--	--	.56 (1.76)	.50 (1.64)
ADHD	--	--	--	1.35 (3.85)	1.23 (3.43)
IRR	--	--	--	--	-.34 (.71)
LPE	--	--	--	--	-.02 (.98)
IRR*LPE	--	--	--	--	-.003 (1.0)
Res. Var.	11.85*	11.89*	11.67*	9.14**	7.70**
Slope (Week)					
Intercept	--	.38 (1.47)**	.95 (2.59)*	.36 (1.44)**	.38 (1.46)**
Week (Quad.)	--	--	-.14 (.87)	--	--
IRR	--	--	--	--	.03 (1.03)
LPE	--	--	--	--	.001 (1.0)
IRR * LPE	--	--	--	--	-.002 (1.0)
Res. Var.	--	.001	.02	.002	NA$^{\gamma}$
-2LL	496.5	484.4	481.8	462.5	458.6
Δ -2LL	--	12.1**	2.6	19.3$^{\pm}$	3.9

Note. Unstandardized estimates presented first, exponentiated estimates presented in parentheses (); * < .05; ** < .01; *** < .0001; $^{\pm}\chi^2$ Difference test not computed because models C and D are not nested; $^{\gamma}$Residual variance for slope not included because random effect of slope was removed in Model E due to negative estimate.

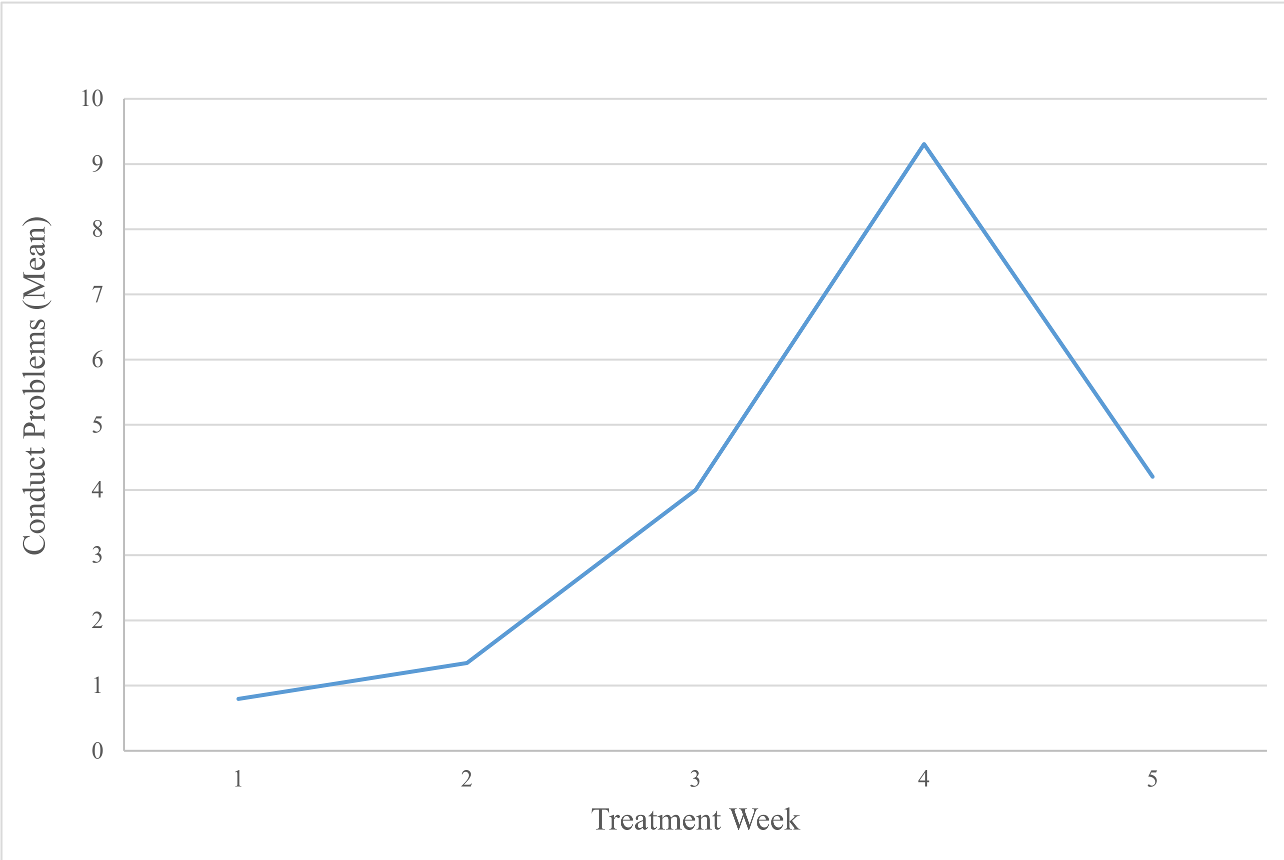

Figure 2. Mean Scores of CP across the STP.

3.3.2 Impairment

Examination of the impairment trajectories across the five treatment weeks was done using PROC MIXED, with results shown in Table 4. First, there was a significant effect of the intercept (b = 2.34, $p < .0001$) in Model A (i.e., unconditional means model), suggesting that week 1 impairment scores were significantly greater than 0. The intraclass correlation for Model A suggested that approximately 70% of the variance in impairment was due to between-person differences. The intercept in Model B was significant (b = 2.45, $p < .0001$), once again suggesting that impairment levels were greater than 0 at week 1. This model included the effect of treatment week (i.e., slope) which was not significant (b = -.05, p = .26).

Next, conditional Model C included the quadratic effect. The G-matrix of this model was not positive definite and led to the removal of the random effect of the quadratic term. The model subsequently converged and found a significant effect of the intercept (b = 2.40, $p < .001$), but no effects of time nor the quadratic effect. Similar to the PROC GLIMMIX model, χ^2 difference tests found that the inclusion of the quadratic effect did not lead to a better fitting model and was not included in the remaining models. Model D, the covariates-only model, found significant effects of the intercept (b = 1.39, $p < .01$) and baseline impairment (b = .36, $p < .01$). Once again, models C and D were compared using AIC/BIC indices given that these were non-nested models. Model C, AIC = 576 and BIC = 589.2 while for Model D, AIC = 542.2 and BIC = 585, suggesting that Model D provided a better fitting model.

Model E included IRR and LPE. Significant effects remained for the intercept and baseline impairment. Further, there was a significant effect of LPE upon the intercept ($b = .03$, $p = .03$), suggesting that higher rates of LPE were associated with significantly higher rates of week 1 impairment. However, like prior models, no significant effects were observed upon the slope of impairment across the five treatment weeks. Further, χ^2 difference tests found that this model was not a significant improvement over Model D ($p = .14$), suggesting that Model D was the best fitting and most parsimonious model.

Table 4. PROC MIXED Effects of Impairment Trajectories.

	Model A	Model B	Model C	Model D	Model E
Initial Status					
Intercept	2.34***	2.45***	2.40***	1.39**	1.41**
Pre-tx IMP	--	--	--	.36**	.33**
Sex (Male)	--	--	--	.26	.33
Medication	--	--	--	.40	.31
Parent Training	--	--	--	.11	.12
ADHD	--	--	--	.25	.06
IRR	--	--	--	--	.04
LPE	--	--	--	--	.03*
IRR*LPE	--	--	--	--	.004
Res. Var.	1.07***	1.27***	1.27***	.70**	.57**
Slope (Week)					
Intercept	--	-.05	.05	-.05	-.04
Week (Quad.)	--	--	-.03	--	--
IRR	--	--	--	--	-.02
LPE	--	--	--	--	.0003
IRR * LPE	--	--	--	--	-.001
Res. Var.	--	.06**	.05**	.06**	.05**
Int/Slope Cov.	--	-.10	-.10	-.05	-.03
L1 Res. Var.	.45***	.32***	.32***	.32***	.32***
-2LL	579.7	563.2	562.0	542.2	532.7
Δ -2LL	--	16.5**	1.2	$19.8^{\pm}$	9.5

Note. * < .05; ** < .01; *** < .0001; $^{\pm}\chi^2$ Difference test not computed because models C and D are not nested.

3.4 Aim 3 – Time-Out

Basic descriptive statistics of time-outs for each child are presented in Table 5. The first set of analyses examined whether IRR and/or LPE predicted changes in the continuous negative behavioral response outcome. Neither Model 1 (covariates-only) nor Model 2 (covariates and predictors) produced any significant effects, suggesting that neither IRR nor LPE (nor any of the covariates) were statistically associated with change in negative behaviors during time-out. The next set of analyses examined whether IRR and/or LPE predicted emotional and/or motivational responses to time-out. After controlling for all theoretically relevant variables, across the 10 outcomes, models found no statistical effects. Thus, results suggest that neither children's emotional response to nor motivation during time-out were contingent upon either IRR or LPE.

Table 5. Descriptive Statistics of TO for each Participant.

			TO Duration	Child-Reported Affect			Counselor-Reported Affect			Counselor-Reported Motivation			
Child	TO *n*	Total # TO	M (SD)	Anger M (SD)	Fun M (SD)	Guilt M (SD)	Anger M (SD)	Fun M (SD)	Guilt M (SD)	Counselors Mad M (SD)	Counselors Laugh M (SD)	Peers Mad M (SD)	Peers Laugh M (SD)
1	1	5	5	2.0	2.0	5.0	2.0	1.0	4.0	1.0	1.0	1.0	1.0
2	1	6	6	1.0	5.0	1.0	1.0	3.0	2.0	1.0	1.0	1.0	1.0
3	21	875	41.67 (41.22)	2.58 (1.92)	1.05 (.30)	1.68 (1.16)	3.39 (1.16)	1.39 (.70)	1.44 (1.25)	3.13 (1.50)	1.22 (.43)	1.61 (.85)	1.39 (1.04)
4	39	1,293	33.15 (21.17)	2.55 (1.39)	1.30 (.85)	2.48 (1.37)	3.36 (1.41)	2.36 (1.27)	1.36 (.74)	3.39 (1.56)	1.76 (1.23)	1.85 (1.25)	2.12 (1.54)
5	13	278	21.38 (21.7)	4.67 (1.16)	1.0	1.33 (1.16)	3.83 (1.03)	1.0	1.67 (.99)	2.0 (1.13)	1.0	1.92 (1.51)	1.0
6	1	5	5	1.0	1.0	6.0	1.0	1.0	3.0	1.0	1.0	1.0	3.0
7	1	7	7	5.0	1.0	5.0	4.0	1.0	4.0	1.0	1.0	1.0	1.0
8	18	380	21.11 (18.31)	2.76 (2.08)	1.82 (1.85)	2.29 (2.17)	2.94 (1.52)	1.47 (.94)	1.59 (.80)	2.0 (1.21)	1.06 (.25)	1.50 (.97)	1.25 (.68)

Notes. TO = Time-out; TO Duration reported in minutes, presented as: Mean (Standard Deviation); Values without a standard deviation indicate a single endorsement on that parameter; For Child-Reported Affect and Counselor-Reported Affect Items, the scale is as follows: 1 = *Not at All*, 2 = *Slightly*, 3 = *Somewhat*, 4 = *Moderately*, 5 = *Extremely*; For Counselor-Reported Motivation Items, the scale is as follows: 1 = *Strongly Disagree*, 2 = *Disagree*, 3 = *Neutral*, 4 = *Agree*, 5 = *Strongly Agree*.

CHAPTER 4. DISCUSSION

IRR and LPE are key factors in understanding the heterogeneity of CP in youth. Yet virtually no work has simultaneously examined the effects of these two traits as it pertains to treatment efficacy. The current study appeared to be the first to investigate whether these distinct etiological mechanisms uniquely predicted treatment efficacy in youth with CP. In the context of an intensive STP, where the majority of youth met diagnostic criteria for oppositional defiant and/or conduct disorder, the current study examined whether IRR and/or LPE were uniquely associated with distinct aspects of treatment efficacy including (1) treatment response, (2) trajectories of CP and impairment change, and (3) response to time-out. Relative to LPE, it was hypothesized that IRR would predict (1) greater treatment response, (2) greater symptom reduction, and (3) greater reductions in behavioral and emotional reactions to time-out. Results showed little support for these hypotheses. One noteworthy finding was that pre-treatment IRR predicted week 5 CP (but not impairment) relative to LPE. Ultimately, however, neither IRR nor LPE were associated with trajectories of symptom change nor response to time-out.

4.1 Treatment Response

The first aim of the study was to examine whether treatment response (i.e., magnitude of change from week 1 to week 5) differed as a function of IRR and/or LPE. It was hypothesized that, after controlling for relevant covariates, IRR would be associated with a greater magnitude of change in week 5 CP and impairment relative to LPE. There was partial support for this hypothesis as pretreatment IRR was significantly and inversely predictive of week 5 CP. Said otherwise, elevated rates of pre-treatment IRR predicted *reduced* week 5 CP. While some prior work has examined how behavioral interventions target emotion dysregulation in youth with disruptive behavior (for review, see Waxmonsky et al., 2021), this study appears to be the first to examine IRR as a *predictor* of change in CP in the context of an intensive behavioral treatment setting. There are a couple of possibilities for this finding. First, youth who presented with greater rates of IRR prior to and at the outset of treatment may have received greater attention from treatment staff, thus reducing their instances of CP over time. It may also be that youth who presented with elevated rates of IRR benefited from the structure of the treatment protocol, which provided opportunities for these youth to improve upon their emotion regulation strategies. These explanations are speculative, as there was no direct measure of whether certain youth received more attention than others. Further, there were no weekly measures of emotion regulation or IRR, so it is unknown whether these constructs changed over time. These results provide future research with a foundation to incorporate measures of underlying treatment mechanisms for more nuanced assessment of treatment efficacy, to determine whether treatments have effectively targeted unique underlying mechanisms of CP.

In contrast to the CP outcome, IRR (and LPE) was unrelated to week 5 impairment. While IRR and LPE show unique associations with impairment (Brotman et al., 2017; Frick et al., 2014b), one possibility for this finding is that caregivers in the current study attributed their child's impairment to disruptive behavior symptoms rather than to IRR (or LPE) specifically (e.g., "my child has a tough time getting along with peers because he interrupts a lot"). Indeed, the IRS assesses for impairment related to specific symptoms rather than impairment due to emotion dysregulation (IRR) or empathy (LPE). It may be that measures assessing impairment due specifically to IRR or LPE would be more sensitive to these affective traits.

LPE, on the other hand, showed no effects upon week 5 CP. This finding adds to the mixed results that exist within the treatment literature. Some prior work has shown that behavior therapy – including the STP – has contributed to greater treatment response, particularly for youth with elevated rates of LPE (Bansal et al., 2019; Hawes et al., 2014; Wilkinson et al., 2016). Yet, few studies have examined treatment response in the same way as measured in the current study. For instance, in contrast to the current findings, Bansal and colleagues (2019) found that higher levels of pre-treatment CP and LPE were associated with the greatest treatment response. However, their analytic approach (i.e., CP as a predictor of treatment response), treatment response scale (i.e., pre- to post-treatment), and measure of LPE (i.e., Antisocial Process Screening Device) differed from the methods used in the current investigation.

The current finding is consistent with other treatment studies – conducted across samples and settings – showing no association between LPE and reductions in CP (Blader et al., 2013; Caldwell, 2011; Kolko & Pardini, 2010; Norlander, 2008). There is some evidence that LPE exerts smaller or no effects upon later aggression and antisocial behavior once other factors associated with disruptive behavior disorders are considered (e.g., impulsivity; Loeber et al., 2009). Another reason for the lack of association may be the broad outcome measure of CP. Treatment staff in the current study coded occurrences of aggressive behaviors (yes vs. no). However, no distinctions were made as to whether the aggressive act was reactive or proactive. This level of nuance in the observation recording procedures would have been particularly insightful since IRR and LPE show unique associations with distinct forms of aggression, with IRR associated with reactive or impulsive forms of aggression and LPE associated with instrumental or planned forms of aggression (Brotman et al., 2017; Frick et al., 2014b; Frick & Marsee, 2006).

4.2 Trajectories of Change

The second aim of the study was to examine whether IRR and LPE predicted trajectories of CP symptom and impairment change across treatment. It was hypothesized that, after controlling for respective baseline measures, IRR would predict a curvilinear change in CP and impairment, whereas LPE would predict a flatter change in CP and impairment. Overall, there was little support for these hypotheses across the MIXED and GLIMMIX models.

LPE predicted higher baseline levels of impairment at week 1 (i.e., the intercept) within the MIXED models. Prior work has shown LPE predicting higher rates of baseline measures but no longitudinal predictive ability; however, these prior studies tested the utility of the DSM-5 LPE specifier (rather than a continuous measure) using early childhood samples in non-treatment settings (Bansal et al., 2021; Déry et al., 2019). This result does align with a large body of work showing that elevated rates of LPE are associated with significantly greater concurrent functional impairment (Frick et al., 2014b, 2014a). Within the GLIMMIX models, there was also an effect of sex such that males presented with significantly higher rates of CP relative to females. This is an interesting finding considering the CP *gender paradox* (e.g., females show lower rates of CP but greater severity; Tiet et al., 2001; Waschbusch, 2002). This finding may be a function of sample size or is sample specific since there were twice as many males (n = 35) as there were females (n = 14), but such an assertion is speculative and warrants further investigation.

In contrast to effects upon the intercept, neither IRR nor LPE predicted changes in slope, suggesting that the trajectory of CP and impairment change did not differ as a function of affect. These findings counter those by Lindhiem and Kolko (2010, 2011) who found that CP symptoms followed a curvilinear decrease over the course of treatment. Key differences between these two studies include treatment type (i.e., cognitive-behavioral classes vs. intensive behavioral program), the measures used to assess CP (i.e., seven-item measure of CP vs. observational ratings), and the fact that prior work did not examine whether certain factors, such as affect, predicted changes in these trajectories. The lack of longitudinal findings in the current study may be due to a couple of reasons. One possibility is that affect only exhibits a small to moderate effect upon later CP. The sample size (see *Limitations*) was likely underpowered for either affective trait (along with their interaction) to exert any kind of statistical effect upon the CP outcome. Another possibility for the lack of trajectory findings is the way that CP were operationalized. More specifically, the behaviors that comprised the CP outcome (i.e., intentional aggression, property destruction, lying, and stealing) only included symptoms consistent with conduct disorder (American Psychiatric Association, 2013). In contrast, this CP outcome did not include symptoms consistent with opposition defiant disorder, despite this disorder being included within the *theoretical* conceptualization of CP (Frick & Morris, 2004). Further, the antisocial behaviors comprising the CP outcome occurred infrequently throughout treatment. For instance, of the 49 youth, 45 (91.8%) did not engage in any intentionally aggressive behaviors during the first week of camp, with similar rates observed for intentional destruction of property. Further, only a single child (2%) was caught lying, with similar rates being observed for stealing. It may be that the base rate of these overt aggressive behaviors was too infrequent (or that the covert behaviors were not observed by the staff) for either affective trait to robustly predict.

One consistent finding that was observed across the GLIMMIX models examining CP change was the effect of the slope. More specifically, CP appeared to *increase* over the course of treatment, and this effect of time was not contingent upon the role of affect. The direction of this result was unanticipated. One possible explanation for this trend is that the gradual uptick in scores was influenced by a select few youth who exhibited increasingly high rates of CP as treatment progressed. For instance, the majority of the sample never engaged in a CP across the five weeks of treatment (71.4% - 85.7% of the sample never exhibiting a CP). The greatest increase in scores occurred in week 4, when two children (4% of the total sample) exhibited 83 and 336 instances of CP. It may be that as treatment went on, these select few children became more frustrated with the protocol and engaged in higher rates of CP.

4.3 Time-Out

The final aim of the study examined whether reactions to time-out differed as a function of affect. It was hypothesized that IRR would be associated with a decrease in negative behaviors and with anger or sadness during time-out, whereas LPE would be associated with an increase in negative behaviors and with amusement during time-out. The small sample size and unbalanced nature of the time-out data precluded within-person analysis and focused solely on between-person effects. Overall, results did not support this hypothesis, as neither IRR nor LPE were associated with changes in total negative behaviors, nor were these affective traits predictive of emotional nor motivational reactions to time-out. These findings counter a handful of studies that examined how time-out is impacted by affect, specifically LPE (Bansal, Haas, et al., 2020; Dadds & Rhodes, 2008; Haas et al., 2011; Hawes & Dadds, 2005). The most likely explanation for the lack of findings is the fact that only eight youth were included in these analyses. This small sample size severely underpowered any type of robust statistical analysis and these findings should be interpreted with extreme caution. The lack of findings is likely better explained by the sample size rather than due to theoretical underpinnings of IRR and/or LPE.

Future research could take certain routes to examine time-out data. Perhaps the most feasible is to accumulate several years' worth of data to conduct somewhat more robust analysis of how affect (or other factors) predict behavioral and emotional reactions to time-out. Alternative statistical methods could have been employed to obtain a more nuanced analysis of the time-out data, such as N-of-1 analyses. This statistical approach allows for personalized analysis at the individual level to determine how time-varying factors predict a given response (Vieira et al., 2017). Applying this approach to time-out, each child would have been analyzed separately to see how their time-out reactions differed as a function of a given variable (e.g., IRR, LPE). N-of-1 analyses require a specific design (i.e., time-varying predictors rather than fixed effects). Time-outs in the current study did not follow that design, yet future research could attempt to design time-outs in the STP to fit this unique analytic approach.

4.4 Limitations

The current study appears to be the first to simultaneously examine how the affective traits of IRR and LPE, the two affective pathways to CP, predict unique aspects of treatment efficacy. The findings from this study offer a foundation for future treatment studies to build upon. With that said, the findings of this study need to be considered in light of several limitations. Perhaps the most salient limitation is the sample size of 49 youth, which is smaller than past treatment studies focusing on LPE (Bansal et al., 2019; Haas et al., 2011; McDonald et al., 2011), IRR/emotion dysregulation (see Waxmonsky et al., 2021), and general disruptive behavior disorders (Pelham Jr. et al., 2000). This is particularly evident for the time-out analysis, which only yielded eight youth across the 2019 STP who received a time-out. The original full sample was intended to be approximately 70 youth between the ages of 7 and 12. This sample size was adequate to obtain power of .80 for a medium effect, $d = .25$ (Faul et al., 2007). However, due to unavoidable circumstances (i.e., global pandemic), this target N was not achieved, yielding a sample of just under 50 youth. Post-hoc power analysis found that the current sample was adequately powered to detect large effects (.40). Offsetting concerns with the sample size is the use of multiple time-points and within-subjects design which appeared adequate for longitudinal multilevel modeling techniques (Maas & Hox, 2005). The sample was relatively homogenous in terms of racial and ethnic diversity. While it was reflective of the community from which it was recruited, these results likely do not generalize to communities with more diverse children. Thus, future research should attempt to recruit larger and more diverse samples that would be large enough to detect smaller effects (e.g., effects of affect) and generalize to youth of various demographic backgrounds.

Next, certain constructs examined in this study likely presented with a low base rate, specifically CP. As aforementioned, the CP outcome variable only captured behaviors consistent with conduct disorder and did not include symptoms of oppositional defiant disorder. The low base rate of those antisocial behaviors likely created a low statistical ceiling of little variability in those responses. This low-risk, community sample did not show enough severity in their antisocial behavior that would have been detected in statistical analyses. In contrast, other behaviors that may have shown greater variation in this sample include interruption and noncompliance, as these are more closely related to ADHD and oppositional defiant disorder, respectively, which was the primary diagnostic make-up of the sample. Indeed, the STP appears most appropriate in targeting disruptive behavior associated with ADHD and oppositional defiant disorder relative to highly severe conduct disorder behaviors (e.g., cruelty to people/animals, use of a weapon, vandalism). Further, and as mentioned previously, the CP outcome was broad and did not discriminate between reactive and proactive forms of aggression. This level of analysis is necessary given that IRR is more closely related to reactive aggression (Brotman et al., 2017; Leibenluft & Stoddard, 2013) whereas LPE is most closely associated with proactive aggression (Frick et al., 2014b; Frick & Marsee, 2006).

The scale of treatment response was limited to *within* treatment change rather than pre- to post-treatment. This change in scale was done given that measures were not uniform at pre- and post-treatment. Future studies should ensure that the same battery of ratings is used at both timepoints to accurately examine treatment response as seen in prior work (Bansal et al., 2019; Lindhiem et al., 2012). Relatedly, the current study was unable to examine treatment *outcomes* (i.e., likelihood of normalization at post-treatment) which is done using a group of typically developing youths. The current study only had three youth who did not meet any DSM-5 diagnosis. This precluded the current study's ability to reliably determine how IRR and LPE were associated with normalization at post-treatment. Therefore, future work should recruit additional typically developing children to analyze how the affective traits of IRR and LPE predict treatment outcomes.

Lastly, the post-time-out items regarding child affect and motivation were designed by a doctoral level clinician and doctoral graduate student. The validity and reliability of these items is unknown and may have contributed to the lack of findings in time-out analyses. Establishing the psychometrics of this measure was beyond the scope of the current study, but future work should examine the psychometric utility of this measure. Such work would offer clinicians several easy-to-administer items to assess a child's affect post-time-out, especially in an intensive and dynamic treatment setting.

4.5 Conclusions

IRR and LPE are important factors in understanding the heterogeneity of CP, and the current study took the first step in simultaneously examining how these affective traits predict treatment efficacy for youth with CP. Overall, there was little support for the hypothesized effects. The one finding that was consistent with hypotheses was that IRR was predictive of week 5 CP treatment response (i.e., a greater magnitude of change). Since there was no analysis of how the different treatment components interacted with affect, it cannot be claimed that treatment was effective for youth with higher pre-treatment IRR. Yet, it can be suggested that youth who present with elevated rates of pre-treatment IRR may experience a reduction in their CP from week 1 to week 5 as part of an intensive summer treatment protocol.

Counter to hypotheses, however, was that neither IRR nor LPE predicted treatment response for impairment, nor did these affective traits show any association with trajectories of CP or impairment change, nor did they predict behavioral or emotional reactions to time-out. It may be inappropriate to wholly conclude that IRR and LPE are not important factors in understanding treatment efficacy. As previously discussed, the lack of findings for these aims may likely be a function of the study methodology, particularly the small and underpowered sample size along with low base rates of LPE and CP. Future research should improve upon these methods by recruiting a larger and more diverse sample of youth, and overrecruiting for youth with higher levels of IRR, LPE, and CP. Future work should also implement repeated measures of underlying treatment mechanisms (e.g., emotion dysregulation and empathy) to determine whether treatment protocols are adequately designed to target the maintenance factors of aggressive behaviors. Other variables of interest should also be included in future analysis, such as caregivers' use of management skills or peer influences.

Ultimately, the findings of the current study provide a foundation for future research to build upon as well as incorporating nuanced measures of underlying treatment mechanisms. Such measures will help move broad, nomothetic treatments towards an era of personalized and idiographic interventions designed to target the latent mechanisms that are contributing to and maintaining aggressive and disruptive CP that are unique to each youth presenting for treatment.

APPENDIX

APPENDIX – TIME-OUT LOG

Child: **Group**: **Date**: **Day#**: **Staff:**

Reason for TO (circle one): IA IDP RNC **TO Process, Full (Possible Reduction)**: 10 (5); 20 (10); 30 (15); 30 max

Time Assigned	Time Escalated	Time Escalated	Time Escalated	Time Started (if delayed)*	Suspended	Out	In	End Time	Extra Mins	PTOSO	Total Min TO	Time Left Inquiries

*Time Started is ONLY required if child's TO was fully escalated but refused to go to the TO area right away.

Physical Management

Start	End	Start	End	Start	End	Start	End	Total

	Total
Intentional Aggression Peer	
Unintentional Aggression Peer	
Intentional Aggression Staff	
Unintentional Aggression Staff	
Intentional Destruction	
Unintentional Destruction	
Noncompliance / Repeated NC	
Stealing	
Leaving Area	
Lying	
Verbal Abuse	
Name Calling / Teasing	
Cursing / Swearing	
Interruption	
Complaining / Whining	
Violating Activity Rules	

NOTES

At the end of TO, **ask the child**:

	Not at All	Slightly	Somewhat	Moderately	Extremely	
How angry did you feel during TO?	1	2	3	4	5	NA
How much fun was TO?	1	2	3	4	5	NA
How guilty did you feel during TO?	1	2	3	4	5	NA

After the TO is over, the **Counselor who monitored the TO** should rate each of the following:

	Not at All	Slightly	Somewhat	Moderately	Extremely	
How angry did the child appear during TO?	1	2	3	4	5	NA
How much fun did the child appear to have during TO?	1	2	3	4	5	NA
How guilty did the child appear during TO?	1	2	3	4	5	NA

After TO is over, the **Counselor who monitored the TO** should rate how much they agree with the following:

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
The child was trying to make their counselor(s) mad	1	2	3	4	5	NA
The child was trying to make their counselor(s) laugh	1	2	3	4	5	NA
The child was trying to make their peer(s) mad	1	2	3	4	5	NA
The child was trying to make their peer(s) laugh	1	2	3	4	5	NA

References

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *Bmj*, *332*(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders. In *DSM-5* (5th ed.). American Psychiatric Association.

Bansal, P. S., Babinski, D. E., Waxmonsky, J. G., & Waschbusch, D. A. (2020). Psychometric properties of parent ratings on the Inventory of Callous-Unemotional Traits in a nationally representative sample of 5-to 12-year-olds. *Assessment*. https://doi.org/10.1177/1073191120964562

Bansal, P. S., Goh, P. K., Eng, A. G., Elkins, A. R., Thaxton, M., & Martel, M. M. (2021). Utility of the limited prosocial emotions specifier in preschoolers with conduct problems. *Assessment*, *Advanced Online Publication*, 1–13. https://doi.org/10.1177/10731911211051070

Bansal, P. S., Haas, S. M., Willoughby, M. T., Coles, E. K., Pelham, W. E., Jr., & Waschbusch, D. A. (2020). A pilot study of emotional response to time-out in children with conduct problems and callous-unemotional traits. *Psychological Reports*, *123*(5), 2017–2037. https://doi.org/10.1177/0033294119884014

Bansal, P. S., Waschbusch, D. A., Haas, S. M., Babinski, D. E., King, S., Andrade, B. T., & Willoughby, M. T. (2019). Effects of intensive behavioral treatment for children with varying levels of conduct problems and callous-unemotional traits. *Behavior Therapy*, *50*(1), 1–14. https://doi.org/10.1016/j.beth.2018.03.003

Blader, J. C., Pliszka, S. R., Kafantaris, V., Foley, C. A., Crowell, J. A., Carlson, G. A., Sauder, C. L., Margulies, D. M., Sinha, C., Sverd, J., Matthews, T. L., Bailey, B. Y., & Davis, W. B. (2013). Callous-unemotional traits, proactive aggression, and treatment outcomes of aggressive children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(12), 1281–1293. https://doi.org/10.1016/j.jaac.2013.08.024

Blair, R. J. R. (1999). Responsiveness to distress cues in the child with psychopathic tendencies. *Personality and Individual Differences*, *27*(1), 135–145. https://doi.org/10.1016/S0191-8869(98)00231-1

Brotman, M. A., Kircanski, K., & Leibenluft, E. (2017). Irritability in children and adolescents. *Annual Review of Clinical Psychology*, *13*, 317–341. https://doi.org/10.1146/annurev-clinpsy-032816-044941

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Byrd, A. L., Hawes, S. W., Burke, J. D., Loeber, R., & Pardini, D. A. (2018). Boys with conduct problems and callous-unemotional traits: Neural response to reward and punishment and associations with treatment response. *Developmental Cognitive Neuroscience*, *30*, 51–59. https://doi.org/10.1016/j.dcn.2017.12.004

Byrd, A. L., Hawes, S. W., Loeber, R., & Pardini, D. A. (2018). Interpersonal callousness from childhood to adolescence: Developmental trajectories and early risk factors. *Journal of Clinical Child & Adolescent Psychology*, *47*(3), 467–482. https://doi.org/10.1080/15374416.2016.1144190

Byrd, A. L., Loeber, R., & Pardini, D. A. (2014). Antisocial behavior, psychopathic features and abnormalities in reward and punishment processing in youth. *Clinical Child and Family Psychology Review*, *17*(2), 125–156. https://doi.org/10.1007/s10567-013-0159-6

Caldwell, M. F. (2011). Treatment-related changes in behavioral outcomes of psychopathy facets in adolescent offenders. *Law and Human Behavior*, *35*(4), 275–287. https://doi.org/10.1007/s10979-010-9239-z

Chorpita, B. F., & Daleiden, E. L. (2009). *CAMHD biennial report: Effective psychosocial interventions for youth with behavioral and emotional needs*. Hawaii Department of Health, Child and Adolescent Mental Health Division.

Christian, R. E., Frick, P. J., Hill, N. L., Tyler, L., & Frazer, D. R. (1997). Psychopathy and conduct problems in children: II. Implications for subtyping children with conduct problems. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(2), 233–241. https://doi.org/10.1097/00004583-199702000-00014

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*(4), 597–600. https://doi.org/10.1017/S0954579400007318

Craig, S. G., & Moretti, M. M. (2018). Profiles of primary and secondary callous-unemotional features in youth: The role of emotion regulation. *Development and Psychopathology*, *31*(4), 1489–1500. https://doi.org/10.1017/S0954579418001062

Dadds, M. R., & Rhodes, T. (2008). Aggression in young children with concurrent callous–unemotional traits: Can the neurosciences inform progress and innovation in treatment approaches? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *363*(1503), 2567–2576. https://doi.org/10.1098/rstb.2008.0029

Déry, M., Bégin, V., Toupin, J., & Temcheff, C. (2019). Clinical utility of the limited prosocial emotions specifier in the childhood-onset subtype of conduct disorder. *The Canadian Journal of Psychiatry*, *64*(12), 838–845. https://doi.org/10.1177/0706743719885469

Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology*, *53*(6), 1146. https://doi.org/10.1037/0022-3514.53.6.1146

Eisenberg, N., Cumberland, A., Spinrad, T. L., Fabes, R. A., Shepard, S. A., Reiser, M., Murphy, B. C., Losoya, S. H., & Guthrie, I. K. (2001). The relations of regulation and emotionally to children's externalizing and internalizing problem behavior. *Child Development*, *72*(4), 1112–1134. https://doi.org/10.1111/1467-8624.00337

Ellis, M. L., Weiss, B., & Lochman, J. E. (2009). Executive functions in children: Associations with aggressive behavior and appraisal processing. *Journal of Abnormal Child Psychology*, *37*(7), 945–956. https://doi.org/10.1007/s10802-009-9321-5

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology*, *37*(1), 215–237. https://doi.org/10.1080/15374410701820117

Fabiano, G. A., Pelham Jr., W. E., Manos, M. J., Gnagy, E. M., Chronis, A. M., Onyango, A. N., Lopez-Williams, A., Burrows-MacLean, L., Coles, E. K., Meichenbaum, D. L., Caserta, D. A., & Swain, S. (2004). An evaluation of three time-out procedures for children with attention-deficit/hyperactivity disorder. *Behavior Therapy*, *35*(3), 449–469. https://doi.org/10.1016/S0005-7894(04)80027-3

Fabiano, G. A., Pelham Jr., W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., Onyango, A. N., Kipp, H., Lopez-Williams, A., & MacLean-Burrows, L. (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*(3), 369–385. https://doi.org/10.1207/s15374424jccp3503_3

Fairchild, A. J., & MacKinnon, D. P. (2009). A General Model for Testing Mediation and Moderation Effects. *Prevention Science*, *10*(2), 87–99. https://doi.org/10.1007/s11121-008-0109-6

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Foster, E. M., Jones, D. E., & The Conduct Problems Prevention Research Group. (2005). The high costs of aggression: Public expenditures resulting from conduct disorder. *American Journal of Public Health*, *95*(10), 1767–1772. https://doi.org/10.2105/AJPH.2004.061424

Frick, P. J. (2001). Effective interventions for children and adolescents with conduct disorder. *The Canadian Journal of Psychiatry*, *46*(7), 597–608.

Frick, P. J. (2004). *The Inventory of Callous-Unemotional Traits*. Unpublished rating scale.

Frick, P. J. (2009). Extending the construct of psychopathy to youth: Implications for understanding, diagnosing, and treating antisocial children and adolescents. *The Canadian Journal of Psychiatry*, *54*(12), 803–812. https://doi.org/10.1177/070674370905401203

Frick, P. J., Cornell, A. H., Barry, C. T., Bodin, S. D., & Dane, H. E. (2003). Callous-unemotional traits and conduct problems in the prediction of conduct problem severity, aggression, and self-report of delinquency. *Journal of Abnormal Child Psychology*, *31*(4), 457–470. https://doi.org/10.1023/A:1023899703866

Frick, P. J., Lilienfeld, S. O., Ellis, M., Loney, B., & Silverthorn, S. (1999). The association between anxiety and psychopathy dimensions in children. *Journal of Abnormal Child Psychology*, *27*(5), 383–392. https://doi.org/10.1023/A:1021928018403

Frick, P. J., & Marsee, M. A. (2006). Psychopathy and the developmental pathways to antisocial behavior in youth. In C. J. Patrick (Ed.), *The Handbook of Psychopathy* (pp. 353–375). Guilford Press.

Frick, P. J., & Morris, A. S. (2004). Temperament and developmental pathways to conduct problems. *Journal of Clinical Child and Adolescent Psychology*, *33*(1), 54–68. https://doi.org/10.1207/S15374424JCCP3301_6

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014a). Annual research review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. *Journal of Child Psychology and Psychiatry*, *55*(6), 532–548. https://doi.org/10.1111/jcpp.12152

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014b). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin*, *140*(1), 1–57. https://doi.org/10.1037/a0033076

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin*, *118*(3), 392–404. https://doi.org/10.1037/0033-2909.118.3.392

Ghandour, R. M., Sherman, L. J., Vladutiu, C. J., Ali, M. M., Lynch, S. E., Bitsko, R. H., & Blumberg, S. J. (2019). Prevalence and treatment of depression, anxiety, and conduct problems in US children. *The Journal of Pediatrics*, *206*, 256-267.e3. https://doi.org/10.1016/j.jpeds.2018.09.021

Graziano, P. A., Slavec, J., Hart, K., Garcia, A., & Pelham, W. E. (2014). Improving school readiness in preschoolers with behavior problems: Results from a summer treatment program. *Journal of Psychopathology and Behavioral Assessment*, *36*(4), 555-569. https://doi.org/10.1007/s10862-014-9418-1

Haas, S. M., Waschbusch, D. A., Pelham Jr., W. E., King, S., Andrade, B. F., & Carrey, N. J. (2011). Treatment response in CP/ADHD children with callous/unemotional traits. *Journal of Abnormal Child Psychology*, *39*(4), 541–552. https://doi.org/10.1007/s10802-010-9480-4

Hawes, D. J., & Dadds, M. R. (2005). The treatment of conduct problems in children with callous-unemotional traits. *Journal of Consulting and Clinical Psychology*, *73*(4), 737. https://doi.org/10.1037/0022-006X.73.4.737

Hawes, D. J., Price, M. J., & Dadds, M. R. (2014). Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: A comprehensive review. *Clinical Child and Family Psychology Review*, *17*(3), 248–267. https://doi.org/10.1007/s10567-014-0167-1

Herpers, P., Scheepers, F. E., Bons, D. M. A., Buitelaar, J. K., & Rommelse, N. N. J. (2014). The cognitive and neural correlates of psychopathy and especially callous–unemotional traits in youths: A systematic review of the evidence. *Development and Psychopathology*, *26*, 1–29. https://doi.org/10.1017/S0954579413000527

Huang, F. L., & Cornell, D. G. (2012). Pick Your Poisson: A tutorial on analyzing counts of student victimization data. *Journal of School Violence*, *11*(3), 187–206. https://doi.org/10.1080/15388220.2012.682010

Hudziak, J. J., Achenbach, T. M., Althoff, R. R., & Pine, D. S. (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research*, *16*(Suppl1), S16-23. https://doi.org/10.1002/mpr.217

Jaffee, S. R., Lombardi, C. M., & Coley, R. L. (2013). Using complementary methods to test whether marriage limits men's antisocial behavior. *Development and Psychopathology*, *25*(1), 65–77. https://doi.org/10.1017/S0954579412000909

Kircanski, K., Clayton, M. E., Leibenluft, E., & Brotman, M. A. (2018). Psychosocial treatment of irritability in youth. *Current Treatment Options in Psychiatry*, *5*(1), 129–140. https://doi.org/10.1007/s40501-018-0141-5

Kircanski, K., Zhang, S., Stringaris, A., Wiggins, J. L., Towbin, K. E., Pine, D. S., Leibenluft, E., & Brotman, M. A. (2017). Empirically derived patterns of psychiatric symptoms in youth: A latent profile analysis. *Journal of Affective Disorders*, *216*, 109–116. https://doi.org/10.1016/j.jad.2016.09.016

Kolko, D. J., Dorn, L. D., Bukstein, O. G., Pardini, D. A., Holden, E. A., & Hart, J. (2009). Community vs. Clinic-based modular treatment of children with early-onset ODD or CD: A clinical trial with 3-year follow-up. *Journal of Abnormal Child Psychology*, *37*(5), 591–609. https://doi.org/10.1007/s10802-009-9303-7

Kolko, D. J., & Pardini, D. A. (2010). ODD dimensions, ADHD, and callous–unemotional traits as predictors of treatment response in children with disruptive behavior disorders. *Journal of Abnormal Psychology*, *119*(4), 713. https://doi.org/10.1037/a0020910

Leibenluft, E., & Stoddard, J. (2013). The developmental psychopathology of irritability. *Development and Psychopathology*, *25*(4pt2), 1473–1487. https://doi.org/10.1017/S0954579413000722

Lindhiem, O., & Kolko, D. J. (2010). Trajectories of symptom reduction and engagement during treatment for childhood behavior disorders: Differences across settings. *Journal of Abnormal Child Psychology*, *38*(7), 995–1005. https://doi.org/10.1007/s10802-010-9416-z

Lindhiem, O., & Kolko, D. J. (2011). Trajectories of symptom reduction during treatment for behavior problems in pediatric primary-care settings. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*(6), 486–494. https://doi.org/10.1007/s10488-011-0335-2

Lindhiem, O., Kolko, D. J., & Cheng, Y. (2012). Predicting psychotherapy benefit: A probabilistic and individualized approach. *Behavior Therapy*, *43*(2), 381–392. https://doi.org/10.1016/j.beth.2011.08.004

Loeber, R., Burke, J. D., Lahey, B. B., Winters, A., & Zera, M. (2000). Oppositional defiant and conduct disorder: A review of the past 10 years, part I. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(12), 1468–1484. https://doi.org/10.1097/00004583-200012000-00007

Loeber, R., Burke, J. D., & Pardini, D. A. (2009). Development and etiology of disruptive and delinquent behavior. *Annual Review of Clinical Psychology*, *5*, 291–310. https://doi.org/10.1146/annurev.clinpsy.032408.153631

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

McDonald, R., Dodson, M. C., Rosenfield, D., & Jouriles, E. N. (2011). Effects of a parenting intervention on features of psychopathy in children. *Journal of Abnormal Child Psychology*, *39*(7), 1013–1023. https://doi.org/10.1007/s10802-011-9512-8

McMahon, R. J., & Frick, P. J. (2010). Evidence-based assessment of conduct problems in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 477–505. https://doi.org/10.1207/s15374424jccp3403_6

McNeish, D. (2017). Small sample methods for multilevel modeling: A colloquial elucidation of REML and the Kenward-Roger correction. *Multivariate Behavioral Research*, *52*(5), 661–670. https://doi.org/10.1080/00273171.2017.1344538

Miller, N. V., Haas, S. M., Waschbusch, D. A., Willoughby, M. T., Helseth, S. A., Crum, K. I., Coles, E. K., & Pelham Jr., W. E. (2014). Behavior therapy and callous-unemotional traits: Effects of a pilot study examining modified behavioral contingencies on child behavior. *Behavior Therapy*, *45*(5), 606–618. https://doi.org/10.1016/j.beth.2013.10.006

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology*, *8*(2), 399–424. https://doi.org/10.1017/S0954579400007161

Monuteaux, M. C., Blacker, D., Biederman, J., Fitzmaurice, G., & Buka, S. L. (2006). Maternal smoking during pregnancy and offspring overt and covert conduct problems: A longitudinal study. *The Journal of Child Psychology and Psychiatry*, *47*(9), 883–890. https://doi.org/10.1111/j.1469-7610.2005.01566.x

Nock, M. K., Kazdin, A. E., Hiripi, E., & Kessler, R. C. (2006). Prevalence, subtypes, and correlates of DSM-IV conduct disorder in the national comorbidity survey replication. *Psychological Medicine*, *36*(5), 699–710. https://doi.org/10.1017/s0033291706007082

Nock, M. K., Kazdin, A. E., Hiripi, E., & Kessler, R. C. (2007). Lifetime prevalence, correlates, and persistence of oppositional defiant disorder: Results from the National Comorbidity Survey Replication. *Journal of Child Psychology and Psychiatry*, *48*(7), 703–713. https://doi.org/10.1111/j.1469-7610.2007.01733.x

Norlander, B. J. (2008). *Targeting dimensions of psychopathy in at -risk youth: Assessment and utility of a focused cognitive behavioral therapy program* [Ph.D., University of North Texas]. https://www.proquest.com/docview/304538263/abstract/DBADFBD4A02C4558PQ/1

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., Poulton, R., Sears, M. R., Thomson, W. M., & Caspi, A. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology*, *20*(2), 673–716. https://doi.org/10.1017/S0954579408000333

Pardini, D. A., & Fite, P. J. (2010). Symptoms of conduct disorder, oppositional defiant disorder, attention-deficit/hyperactivity disorder, and callous-unemotional traits as unique predictors of psychosocial maladjustment in boys: Advancing an evidence base for DSM-V. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(11), 1134–1144. https://doi.org/10.1016/j.jaac.2010.07.010

Pardini, D. A., Lochman, J. E., & Frick, P. J. (2003). Callous/unemotional traits and social-cognitive processes in adjudicated youths. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(3), 364–371. https://doi.org/10.1097/00004583-200303000-00018

Pelham Jr., W. E., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 449–476. https://doi.org/10.1207/s15374424jccp3403_5

Pelham Jr., 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 and Adolescent Psychiatry*, *31*(2), 210–218. https://doi.org/10.1097/00004583-199203000-00006

Pelham Jr., W. E., Gnagy, E. M., Greiner, A. R., Fabiano, G. A., Waschbusch, D. A., & Burrows-MacLean, L. (2010). Summer treatment programs for attention-deficit/hyperactivity disorder. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidenced-based psychotherapies for children and adolescents* (2nd ed., pp. 277–292). Guilford Press.

Pelham Jr., W. E., Gnagy, E. M., Greiner, A. R., Hoza, B., Hinshaw, S. P., Swanson, J. M., Simpson, S., Shapiro, C., Bukstein, O., Baron-Myak, C., & McBurnett, K. (2000). Behavioral versus behavioral and pharmacological treatment in ADHD children attending a summer treatment program. *Journal of Abnormal Child Psychology*, *28*(6), 507–525. https://doi.org/10.1023/A:1005127030251

Pelham Jr., W. E., Greiner, A. R., & Gnagy, E. M. (2017). *Children's summer treatment program manual*. Florida International University, Center for Children and Families.

Pelham Jr., W. E., & Hoza, B. (1996). *Intensive treatment: A summer treatment program for children with ADHD*.

Ray, J. V., & Frick, P. J. (2018). Assessing callous-unemotional using the total score from the Inventory of Callous-Unemotional Traits: A meta-analysis. *Journal of Clinical Child and Adolescent Psychology*, 1–10. https://doi.org/10.1080/15374416.2018.1504297

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems – a meta-analytic review. *Journal of Child Psychology and Psychiatry*, *47*(1), 99–111. https://doi.org/10.1111/j.1469-7610.2005.01544.x

Schabenberger, O. (2007). *Growing Up Fast: SAS 9.2 enhancements to the GLIMMIX procedure*. SAS Institute, Inc.

Stringaris, A., Goodman, R., Ferdinando, S., Razdan, V., Muhrer, E., Leibenleft, E., & Brotman, M. A. (2012). The Affective Reactivity Index: A concise irritability scale for clinical and research settings. *The Journal of Child Psychology and Psychiatry*, *53*(11), 1109–1117. https://doi.org/10.1111/j.1469-7610.2012.02561.x

Thomson, N. D., & Centifanti, L. C. M. (2018). Proactive and reactive aggression subgroups in typically developing children: The role of executive functioning, psychophysiology, and psychopathy. *Child Psychiatry and Human Development*, *49*(2), 197–208. https://doi.org/10.1007/s10578-017-0741-0

Tiet, Q. Q., Wasserman, G. A., Loeber, R., McReynolds, L. S., & Miller, L. S. (2001). Developmental and sex differences in types of conduct problems. *Journal of Child and Family Studies*, *10*(2), 181–197. https://doi.org/10.1023/A:1016637702525

Vieira, R., McDonald, S., Araújo-Soares, V., Sniehotta, F. F., & Henderson, R. (2017). Dynamic modelling of n-of-1 data: Powerful and flexible data analytics applied to individualised studies. *Health Psychology Review*, *11*(3), 222–234. https://doi.org/10.1080/17437199.2017.1343680

Wakschlag, L. S., Estabrook, R., Petitclerc, A., Henry, D., Burns, J. L., Perlman, S. B., Voss, J. L., Pine, D. S., Leibenluft, E., & Briggs-Gowan, M. L. (2015). Clinical implications of a dimensional approach: The normal:abnormal spectrum of early irritability. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(8), 626–634. https://doi.org/10.1016/j.jaac.2015.05.016

Waschbusch, D. A. (2002). A meta-analytic examination of comorbid hyperactive-impulsive-attention problems and conduct problems. *Psychological Bulletin*, *128*(1), 118–150. https://doi.org/10.1037/0033-2909.128.1.118

Waschbusch, D. A., Baweja, R., Babinski, D. E., Mayes, S. D., & Waxmonsky, J. G. (2020). Irritability and limited prosocial emotions/callous-unemotional traits in elementary-school-age children. *Behavior Therapy*, *51*(2), 223–237. https://doi.org/10.1016/j.beth.2019.06.007

Waschbusch, D. A., Carrey, N. J., Willoughby, M. T., King, S., & Andrade, B. F. (2007). Effects of methylphenidate and behavior modification on the social and academic behavior of children with disruptive behavior disorders: The moderating role of callous/unemotional traits. *Journal of Clinical Child and Adolescent Psychology*, *36*(4), 629–644. https://doi.org/10.1080/15374410701662766

Waschbusch, D. A., & Willoughby, M. T. (2008). Attention-deficit/hyperactivity disorder and callous-unemotional traits as moderators of conduct problems when examining impairment and aggression in elementary school children. *Aggressive Behavior*, *34*(2), 139–153. https://doi.org/10.1002/ab.20224

Waxmonsky, J. G., Baweja, R., Bansal, P. S., & Waschbusch, D. A. (2021). A review of the evidence base for psychosocial interventions for the treatment of emotion dysregulation in children and adolescents. In M. P. Singh & G. A. Carlson (Eds.), *Child and Adolescent Psychiatric Clinics of North America* (pp. 573–594). Elsevier. https://www.sciencedirect.com/science/article/pii/S1056499321000237

Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology*, *65*(1), 93–109. https://doi.org/10.1037/0022-006X.65.1.93

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children—Fifth Edition*. Pearson.

Wilkinson, S., Waller, R., & Viding, E. (2016). Practitioner review: Involving young people with callous unemotional traits in treatment—Does it work? A systematic review. *The Journal of Child Psychology and Psychiatry*, *57*(5), 552–565. https://doi.org/10.1111/jcpp.12494

VITA

PEVITR S. BANSAL, M.S.

EDUCATION

Ph.D. **University of Kentucky**, Lexington, KY — In Progress (expected, 2023)
Clinical Psychology
Advisor: Michelle M. Martel, Ph.D.
Dissertation: *The role of affective heterogeneity on treatment effects for youth with conduct problems*

M.S. **University of Kentucky**, Lexington, KY — 2019
Clinical Psychology
Advisor: Michelle M. Martel, Ph.D.
Thesis: *Empirical assessment of callous-unemotional traits in preschool: A comparison of confirmatory factor and network analysis*

M.S. **Philadelphia College of Osteopathic Medicine**, Philadelphia, PA — 2015
Counseling and Clinical Health Psychology
Advisor: Elizabeth A. Gosch, Ph.D., A.B.P.P.

B.A. **Pennsylvania State University**, Abington, PA — 2014
Psychology

RESEARCH EXPERIENCE

University of Kentucky, RISK Lab — Lexington, KY
Department of Psychology, *Graduate Student* — August 2017 – Present
Supervisor: Michelle M. Martel, Ph.D.

Penn State Hershey Medical Center, College of Medicine — Hershey, PA
Department of Psychiatry, *Human Research Technologist* — August 2015 – Present
Supervisor: Daniel A. Waschbusch, Ph.D., A.B.P.P.

Children's Hospital of Philadelphia — Philadelphia, PA
Center for Management of ADHD, *Research Assistant* — January 2014 – August 2015
Supervisors: Jenelle Nissley – Tsiopinis, Ph.D. & Thomas J. Power, Ph.D.

Pennsylvania State University – Abington College — Abington, PA
Department of Psychology, *Research Assistant* — January 2013 – May 2013
Supervisors: Jacob A. Benfield, Ph.D., & Michael J. Bernstein, Ph.D.

CLINICAL EXPERIENCE

University of Kentucky, Good Samaritan Hospital Lexington, KY
Behavioral Health Inpatient Unit, *Practicum Student* August 2021 – Present
Clinical Supervisor: Michelle M. Martel, Ph.D.

University of Kentucky, Psychological Services Center Lexington, KY
Department of Psychology, *Senior Assessment Consultant* August 2021 – Present
Clinical Supervisor: Michelle M. Martel, Ph.D.

The Lexington School Lexington, KY
Graduate Assessment Examiner April 2021 – May 2021
Clinical Supervisor: Mary Beth McGavran, Ph.D.

University of Kentucky, Psychological Services Center Lexington, KY
Department of Psychology, *Assistant Director* July 2020 – July 2021
Clinical Supervisor: David T. Susman, Ph.D.

Mind Psi: Psychological Services for Children and Adolescents Lexington, KY
Practicum Student August 2019 – April 2020
Clinical Supervisor: Michelle M. Martel, Ph.D.

Penn State Hershey, Summer Treatment Program Harrisburg, PA
Department of Psychiatry, *Assistant Director* June 2019 – August 2019
Program Director & Clinical Supervisor: Daniel A. Waschbusch, Ph.D., A.B.P.P.

University of Kentucky, Psychological Services Center Lexington, KY
Department of Psychology, *Graduate Therapist* September 2018 – December, 2021
Clinical Supervisors: Mary Beth McGavran, Ph.D.; Heather Risk, PsyD.; Lindsey Jansinski, Ph.D.

University of Kentucky, Counseling Center Lexington, KY
Graduate Therapist August 2018 – May 2019
Clinical Supervisor: Alyssa Frye, Psy.D., & Walter Malone, M.A.

Penn State Hershey, Summer Treatment Program Harrisburg, PA
Department of Psychiatry, *Graduate Counselor* June 2018 – August 2018
Program Director & Clinical Supervisor: Daniel A. Waschbusch, Ph.D., A.B.P.P.

Penn State Hershey, Psychiatry Outpatient Services Hershey, PA
Department of Psychiatry, *Human Research Technologist* August 2015 – July 2017
Clinical Supervisors: Erika F. H. Saunders, M.D. & Dara E. Babinski, Ph.D.

Empowerment Resources Associates Philadelphia, PA
Graduate Therapist September 2014 – June 2015
Supervisor: Christopher Brown, M.S., L.P.C. & Alan Schwartz, Psy.D.

Florida International University, Summer Treatment Program Miami, FL
Center for Children & Families, *Paraprofessional Counselor* June 2012 – August 2012
Program Director: William E. Pelham, Jr., Ph.D.
Supervisor: Daniel A. Waschbusch, Ph.D., A.B.P.P.

PUBLICATIONS

1. **Bansal, P. S.**, Goh, P. K., Southward, M. W., & Martel, M. M. (Under Review). Impulsivity as key bridge symptoms in cross-sectional and longitudinal networks of ADHD and ODD.

2. Brennan, G., **Bansal, P. S.**, Waxmonsky, J. W., Waschbusch, D. A., & Babinski, D. E. (Under Review). Associations among ADHD symptoms, ODD symptoms, and borderline personality feature: A network analysis.

3. Martel, M. M., Elkins, A. R., Eng, A. G., Goh. P. K., **Bansal, P. S.**, Smith, T. E., Thaxton, M. H., Ryabinin, P., Mooney, M. A. Gustafsson, H. C., Karalunas, S. L., & Nigg, J. T. (Accepted). Longitudinal temperament pathways to ADHD between childhood and adolescence. *Research on Child and Adolescent Psychology*.

4. Goh, P. K., Smith, T. E., Lee, C. A., **Bansal, P. S.**, Eng, A. G., & Martel, M. M. (In Press). Etiological networks of attention-deficit/hyperactivity disorder during childhood and adolescence. *Journal of Clinical Child and Adolescent Psychology*.

5. Goh, P. K., Martel, M. M., Jones, P., **Bansal, P. S.**, Eng, A. G., Elkins, A. R., Thaxton, M., & Barkley, R. A. (2021). Clarifying relations between ADHD and functional impairment in Adulthood: Utilization of network and machine learning approaches. *Assessment*. Advanced online publication. https://doi.org/10.1177/10731911211050921

6. **Bansal, P. S.**, Goh, P. K., Eng, A. G., Elkins, A. R., Thaxton, M. H., & Martel, M. M. (2021). Utility of the limited prosocial emotions specifier in preschoolers with conduct problems. *Assessment*. Advanced online publication. https://doi.org/10.1177/10731911211051070

7. **Bansal, P. S.**, Goh, P. K., Eng, A. G., Elkins, A. R., Thaxton, M., Smith, T. E., & Martel, M. M. (2021). Identifying the inter-domain relations among ODD, CD, and CU traits in preschool children using network analysis. *Research on Child and Adolescent Psychopathology*, 1-13.

8. Martel, M. M., Eng, A. G., **Bansal, P. S.**, Smith, T. E., Elkins, A. R., & Goh, P. K. (2021). Multiple informant average integration of ADHD symptom ratings predictive of concurrent and longitudinal impairment. *Psychological Assessment*, *33*(5), 443-451.

9. **Bansal, P. S.**, Babinski, D. E., Waxmonsky, J. G., & Waschbusch, D. A. (2020). Psychometric properties of parent ratings on the Inventory of Callous-Unemotional Traits in a nationally representative sample of 5 to 12-year-olds. *Assessment*, 1-15.

10. Babinski, D. E., Mills, S. L., **Bansal, P. S.**, Breaux, R. P., & Waschbusch, D. A. (2020). Behavioral treatment for social-emotional difficulties of preadolescent and adolescent girls with ADHD. *Evidence-Based Practice in Child and Adolescent Mental Health*, *5*(2), 173-188.

11. **Bansal, P. S.**, Goh, P. K., Lee, C. A., & Martel, M. M. (2020). Conceptualizations of callous-unemotional traits in preschool through confirmatory factor and network analysis. *Journal of Abnormal Child Psychology*. *48*(4), 1-12.

12. Goh, P. K., Lee, C. A., **Bansal, P. S.**, Aguerrevere, L. E, Rucker, A. T., & Martel, M. M. (2020). Interpretability and external validity of a bifactor model of ADHD in young adults: An assessment of general "g" and specific IA and HI factors. *Journal of Psychopathology and Behavioral Assessment.* 1-15.

13. **Bansal, P. S.**, Haas, S. M., Willoughby, M. T., Coles, E. K., Pelham Jr., W. E., & Waschbusch, D. A. (2020). A pilot study of emotional response to time-out in children with conduct problems and callous-unemotional traits. *Psychological Reports. 123*(5), 2017-2037.

14. **Bansal, P. S.**, Waschbusch, D. A., Haas, S. M., Babinski, D. E., King, S., Andrade, B. F., & Willoughby, M. T. (2019). Effects of intensive behavioral treatment on children with varying levels of conduct problems and callous-unemotional traits. *Behavior Therapy. 50*(1), 1-14.

15. Petrovic-Dovat, L. Kumar, M. G., & **Bansal, P. S.** (2018). Refining training opportunities for pediatric and psychiatric residents and fellows within an integrated healthcare model. *Journal of Psychiatry and Behavioral Sciences, 1006*(2), 1-3.

16. Petrovic-Dovat, L., Berlin Jr., C. M., **Bansal. P. S.**, & Fogel, B. N. (2018). Designing programs and educational opportunities to aid pediatricians in treating children with mental illness in an outpatient academic pediatric setting. *Journal of Psychology and Psychiatry, 2*(1), 1-3.

17. Babinski, D. E., Mills, S. L., & **Bansal, P. S.** (2018). The effects of behavioral parent training with adjunctive social skills training for a preadolescent girl with ADHD and borderline personality features. *Clinical Case Studies*, *17*(1), 21-37.

18. Petrovic-Dovat, L., Fausnight, T., White, A. M., Zeiger, T. S., **Bansal, P. S.**, Garg, N.,... & Bixler, E. O. (2016). Degree of anxiety in food allergic children in a tertiary care center. *Annals of Allergy, Asthma & Immunology*, *116*(6), 528-532.