The Dissertation Committee for Zohra Chahal Certifies that this is the approved version of the following Dissertation:

An Adapted Summer Treatment Program for Children with ADHD: Investigating Program Effectiveness and Moderators of Treatment Outcome

Committee:

Kevin Stark, Co-Supervisor

Cindy Carlson, Co-Supervisor

Jane Gray

Mark Bond

Sarah Tannenbaum

An Adapted Summer Treatment Program for Children with ADHD: Investigating Program Effectiveness and Moderators of Treatment Outcome

by

Zohra Chahal

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2022

Dedication

I would like to dedicate this body of work to my beloved grandmother, mother, and father. You instilled in me the courage to pursue my dreams and provided me with unconditional support when my confidence wavered. As I look back at the number of challenges that arose, reflect on where I am now, and look toward the future, excitedly, you help me remember that everything unfolds just as it needs to: in the best way.

Abstract

An Adapted Summer Treatment Program for Children with ADHD: Investigating Program Effectiveness and Moderators of Treatment Outcome

Zohra Chahal, PhD

The University of Texas at Austin, 2022

Co-Supervisor: Kevin Stark

Co-Supervisor: Cindy Carlson

Attention-deficit/hyperactivity disorder (ADHD) is a chronic neurodevelopmental disorder. Clinical practice guidelines established by the American Academy of Pediatrics recommend behavioral treatments as a first-line intervention for preschool and elementary-aged children with ADHD. The Summer Treatment Program (STP) is one such treatment, providing intensive intervention to children with ADHD in the form of an 8-week summer day treatment program. Despite promising outcomes, the STP model remains largely cost-prohibitive for mental health teams and for families. Camp Baker, developed by Judge Baker Children's Center, is a 6-week adaptation of the STP model intended to be more feasible and accessible to children and families in need than the traditional 8-week program. Despite the preliminary evidence and support for the implementation of this adapted STP, further investigation is needed to understand the overall effectiveness of this treatment model in terms of clinically meaningful and interpretable outcomes and to identify subpopulations of children with particularly strong or poor response to this specific intervention. This is the first study to investigate both program effectiveness and moderators of treatment outcome for participants of a 6-week adaptation of the STP delivered in a community setting.

Table of Contents

Chapter I: Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a chronic neurodevelopmental disorder affecting approximately 9.4% of children in the United States (Danielson et al., 2018a). Left unaddressed, the long term effects of ADHD are substantial. In regards to educational outcomes, youth with ADHD are more likely to experience academic underachievement, grade retention, expulsion, and dropout compared to same-age peers without ADHD (Barkley et al., 2006; Owens & Hinshaw, 2016). ADHD in childhood is also associated with an increased risk for encounters with police, unintentional injuries, and risky driving in adolescence (Lahey et al., 2016). Children and adolescents with ADHD are also more likely to experience impaired interpersonal functioning, poorer occupational attainment, elevated health risks, and early mortality in adulthood (Barkley et al., 2006; Cortese et al., 2013; Dalsgaard et al., 2015; Faraone, 2015; Mikami et al., 2008; Nigg, 2013). These functional impairments are compounded in the context of a comorbid psychiatric disorder such as oppositional defiant disorder (ODD), conduct disorder (CD), anxiety, and depression (Cuffe et al., 2015). Specific to youth with ADHD, effective and timely intervention during one of two "sensitive periods" in development (i.e., the transition from childhood to adolescence and/or the transition from adolescence to childhood) carries implications for long-term functioning (Turgay et al., 2012). To address this concern in the literature, the sample used for this study consists of children, ages 6 to 12, with ADHD.

Clinical practice guidelines established by the American Academy of Pediatrics (AAP) recommend behavioral treatments as a first-line intervention for preschool and elementary-

aged children with ADHD (Wolraich et al., 2019). The Summer Treatment Program (STP; Pelham & Hoza, 1996) is one such established behavioral treatment, providing intensive intervention to children with ADHD in the form of an 8-week summer day treatment program and concurrent behavioral parent training to caregivers in the form of once weekly group sessions. Successful completion of the STP is associated with significant reductions ADHD symptomatology and high rates of parent-reported treatment acceptability (Pelham & Hoza, 1996; Fabiano et al., 2014). However, the STP model remains largely cost-prohibitive for community mental health teams to provide and for families to engage in. To facilitate increased dissemination and implementation of the gold standard of treatment for children with ADHD, it is important to investigate the effectiveness of adapted STP models which increase implementation feasibility for community mental health centers while providing robust treatment outcomes, similar to those of the standard STP, for children and their families.

Camp Baker, developed by Judge Baker Children's Center, is a 6-week adaptation of the STP and is designed to be more feasible and accessible to children and families in need than the traditional 8-week program. Preliminary analyses suggest Camp Baker leads to reductions in caregiver-reported symptom severity for attention, externalizing, and internalizing problems (Tannenbaum et al., 2019). Furthermore, caregivers who participate in the behavioral parent training sessions report high overall satisfaction with the program and report feeling more effective as caregivers for their child (Tannenbaum et al., 2019). Despite the preliminary evidence and support for the implementation of this adapted STP, further investigation is needed to understand the overall effectiveness of this treatment model in terms of clinically

meaningful and interpretable outcomes and the subpopulations of children with particularly strong or poor response to this specific intervention. To this end, this study seeks to examine behavioral improvement as defined by parent-reported measures of child functioning across six key domains: inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations. Supplementary measures of behavioral improvement over time will also be used to evaluate program effectiveness, including: changes in frequency of positive and negative behaviors, as defined in the standard STP point system, and changes in percentage of daily report card target behaviors achieved. This study also seeks to identify subgroups of children with differential treatment outcomes, namely by investigating whether children who are engaged in concurrent pharmacotherapy experience greater benefits from the 6-week STP.

In sum, this study seeks to contribute to the existing treatment outcome literature by extending upon what is known about the clinical utility of the STP model. To that end, this is the first study to investigate both program effectiveness and concurrent pharmacotherapy as a moderator of treatment outcome for participants of a 6-week adaptation of the STP delivered in a community setting. Study findings may enhance dissemination and greater implementation of the adapted STP in community care settings across the US. Study findings will provide future directions for continued adaptability research designed to make the STP more feasible to deliver, more efficacious for the heterogenous ADHD population in childhood, and more accessible to families in need of services.

Chapter II: Review of the Literature

Overview of Attention-Deficit/Hyperactivity Disorder

Since the turn of the nineteenth century, there has been an evolving discourse regarding the apt characterization of individuals who demonstrate persistent difficulty sustaining attention and atypical hyperactivity or impulsivity (Lange et al., 2010). The second edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM; DSM-II; (American Psychiatric Association, 1968) labeled this syndrome as "Hyperkinetic Reaction of Childhood", emphasizing the presence of abnormal hyperactivity and describing the disorder as one that typically manifests in childhood and remits by adolescence. Subsequent to research findings indicating deficits in attention and difficulty regulating impulses as more salient characteristics of the syndrome (Douglas, 1972), the disorder was renamed "Attention Deficit Disorder (ADD), with or without hyperactivity" in the DSM-III (American Psychiatric Association, 1980). However, this classification was controversial due to the limited research on and evidence for the validity of the two subtypes of ADD (Barkley, 2015). In the revised diagnostic manual (DSM-III-R; (American Psychiatric Association, 1987), the syndrome was renamed "Attention Deficit-Hyperactivity Disorder (ADHD)" whereby manifestations of the syndrome without symptoms of hyperactivity were relegated to the diagnosis of "Undifferentiated ADD". Prior to the release of the next edition of the DSM, a large scale field study provided evidence for three distinct subtypes of ADHD: (1) predominantly inattentive, (2) predominantly hyperactive-impulsive, and (3) combined (Lahey et al., 1994). These findings informed the classification of ADHD in the subsequent DSM (DSM-IV; (American Psychiatric Association, 1994). The DSM-IV was also the

first edition of the DSM to delineate two categories of ADHD symptoms – that of inattention and that of hyperactivity/impulsivity.

The diagnostic criteria and subtypes of ADHD remain unchanged in DSM-5 (American Psychiatric Association, 2013) with the exception of the maximum age of onset of symptoms, a threshold which increased from 7 years of age to 12 years of age. Most notably, however, the DSM-5 is the first to characterize ADHD as a neurodevelopmental disorder, a term used to describe developmental deficits that manifest in early childhood and often persist throughout the lifespan. In support of this form of classification, neuroimaging research has demonstrated structural and functional abnormalities in certain cortical, subcortical, and cerebellar structures and systems of individuals with ADHD (Albajara Sáenz et al., 2019). Furthermore, there is evidence that 60% of children with ADHD exhibit symptoms into adulthood while 41% of children with ADHD demonstrate both symptoms and clinically significant impairment in adulthood (Sibley et al., 2017).

Dual Pathway Model of ADHD

Of the various models developed to explain the constellation of neurological and functional impairments in individuals with ADHD, the dual pathway model (Sonuga-Barke, 2002, 2003) is the one that most comprehensively accounts for the neurophysiological heterogeneity of ADHD. The dual pathway model posits that there are two pathways, established during early development, by which neurophysiological phenomenon manifest as ADHD: one of executive functioning (EF) deficits (Barkley, 1997) and one of altered-reinforcement (Sagvolden et al., 1998).

The executive function deficit hypothesis (Barkley, 1997) represents the first pathway in the dual pathway model. This hypothesis states that deficient inhibitory control, resulting from alterations of fronto-dorsal striatal circuit networks, prevent the executive functions of working memory and self-regulation of affect, motivation, and arousal from functioning optimally (Barkley, 1997; Sonuga-Barke, 2003). Thus, impairments in behavioral inhibition and EF processes lead to poorer motor control, characteristic of individuals with ADHD.

The delay aversion hypothesis (Sagvolden et al., 1998) represents the second pathway in the dual pathway model. This hypothesis states that individuals with ADHD have a distinct motivational style, in which altered fronto-ventral striatal reward networks and meso-limbic neural projections which terminate in the nucleus accumbens are implicated, that is characterized by an aversion to delayed reinforcement (Sagvolden et al., 1998; Sonuga-Barke, 2003). More specifically, the neurodevelopmental deficits within the reward centers of the brain result in a different delay gradient in individuals with ADHD, which then manifests as difficulty tolerating delayed reinforcement (Sonuga-Barke, 2003).

A meta-analysis of 83 studies revealed an absence of any consistent patterns of EF deficits associated with ADHD (Willcutt et al., 2005) suggesting that although EF deficits may result in clinically significant levels of inattention or hyperactivity/impulsivity, these EF deficits are not always sufficient in generating the symptomatology and degree of impairment characteristic of the diagnosis of ADHD. Thus, the executive function deficit hypothesis cannot be used to describe the neurodevelopmental state of every individual with ADHD. In support of the dual-pathway model, research suggests that EF deficits are primarily related to symptoms

of inattention (Thorell, 2007; Willcutt et al., 2005) whereas problems characterized by high affective involvement, such as delay aversion, are primarily related to symptoms of hyperactivity-impulsivity (Castellanos et al., 2006; Thorell, 2007). The absence of significant interaction effects between EF deficits and delay aversion (Thorell, 2007) provides further evidence for the proposal of two distinct pathways to ADHD (Sonuga-Barke, 2002) as neither the executive function deficit hypothesis nor the delay aversion hypothesis is able to independently account for all heterogenous manifestations of ADHD.

Clinical Presentation

Per the DSM-5, a diagnosis of ADHD is warranted when a child persistently experiences symptoms of inattention and/or hyperactivity-impulsivity – meeting a minimum threshold of six of the nine possible symptoms in either of the two subcategories – over a period of six months, resulting in functional impairment in two or more settings (APA, 2013). Upon meeting these criteria, a child is diagnosed with one of three ADHD subtypes: ADHD, predominantly inattentive presentation (ADHD-I) if the child only demonstrates a significant number of inattentive symptoms; ADHD, predominantly hyperactive/impulsive presentation (ADHD-HI) if the child only demonstrates a significant number of hyperactive/impulsive symptoms; or ADHD, combined presentation (ADHD-C) if the child demonstrates a significant number of both inattentive and hyperactive/impulsive symptoms. Regardless of which subtype a child is diagnosed with, children with ADHD commonly experience functional impairment and distress in both academic and social settings (APA, 2013) resulting in lower global ratings of adaptive functioning (Biederman et al., 1996; Lahey et al., 1998). The scope of functional impairments

experienced by children with ADHD in academic and social settings is discussed in further detail below.

Academic Impairments. Children with ADHD have been shown to demonstrate academic skill deficits in the domains of memory and reasoning as early as preschool (Dupaul et al., 2001), suggesting that these children begin their academic careers at a distinct disadvantage compared to their non-ADHD peers. There is also evidence that ADHD attenuates a child's memory and behavioral performance in the classroom thereby limiting the child's ability to participate in and benefit from developmentally appropriate scholastic demands (Rapport et al., 1999). Symptoms of inattention have also been shown to be predictive of poorer academic performance (Diamantopoulou et al., 2007). Of additional clinical importance, executive functioning deficits have been shown to interact with symptoms of inattention to produce academic underachievement in children with ADHD as early as kindergarten (Thorell, 2007).

Findings from a meta-analysis revealed the effect sizes for a measure of intellectual ability (Full Scale IQ; FSIQ) was significantly different between children with ADHD and their non-ADHD peers (weighted $d = .61$) (Frazier et al., 2004). Although no significant differences in FSIQ were found between children of varying ADHD subtypes, these findings further highlight the degree to which children with ADHD are differentiated from their non-ADHD peers in the academic setting.

Children with ADHD are more likely to demonstrate academic underachievement – broadly defined as below average academic performance – compared to their non-ADHD peers. Children with ADHD have been shown to score significantly lower on achievement tests in the

areas of reading and mathematics (Biederman et al., 1996; Diamantopoulou et al., 2007; Lahey et al., 1998; Frazier et al., 2007), even in cases where children with ADHD do not exhibit executive functioning deficits (T. D. Barry et al., 2002). Children with ADHD have also been described as experiencing academic underachievement as measured by lower scores on standardized assessments (Barbaresi et al., 2007a; Biederman et al., 1996) and lower grade point averages (Frazier et al., 2007).

Compared to non-ADHD peers, children with ADHD are also more likely to experience higher rates of: special education service utilization (LeFever et al., 2002), grade retention (Barbaresi et al., 2007a; Biederman et al., 1996; LeFever et al., 2002), suspension (LaFever et al., 2002), absenteeism (Barbaresi et al., 2007a), expulsion (LaFever et al., 2002), and dropout from school (Barbaresi et al., 2007a; Biederman et al., 1996). A greater number of ADHD symptoms has been found to be predictive of greater utilization of special education services (Diamantopoulou et al., 2007). Among children with ADHD, rates of grade retention increase from middle- to high-school (8.6% at age 12 to >20% at age 19) highlighting the need for earlier intervention and ongoing support as the rigor of the academic curriculum increases over time (Barbaresi et al., 2007a). The risks for school absenteeism, grade retention, and dropout is heightened in the presence of learning disorder or other psychiatric comorbidities (Barbaresi et al., 2007b).

In addition to these disparate functional outcomes, children with ADHD are also more likely to be negatively evaluated by their teachers. On normed measures, children with ADHD have been shown to receive poorer teacher ratings of their academic skills and performance and

lower teacher-reported levels of motivation and effort (McConaughy et al., 2011). Per teacher reports, children with ADHD also demonstrate greater rates of problem behaviors in the classroom (Dupaul et al., 2001) and are less cooperative than their peers (Lahey et al., 1998). Additionally, children with ADHD-HI or ADHD-C, two subtypes of ADHD in which children demonstrate a significant number of hyperactive/impulsive symptoms, are more likely to be described by teachers as disruptive in the classroom (Lahey et al., 1998).

Social Impairments. A greater number of ADHD symptoms has been associated with lower levels of social competence, defined by variables such as: asking for help in an appropriate manner, cooperating with peers in a variety of situations, and adjusting to different behavioral expectations in different settings (Merrell & Boelter, 2001). Children with ADHD subsequently demonstrate behavioral deficits associated with social skills, commonly labeled as antisocial behaviors (Merrell & Boelter, 2001). Such behaviors include: inappropriate assertiveness, noncompliance, arguing, and physical aggression (Cervantes et al., 2013; Merrell & Boelter, 2001). Impairments in demonstrated social skills have been observed in children with ADHD as early as preschool (DuPaul et al., 2001). These social skills deficits carry negative implications for children's interpersonal functioning across a range of settings.

Peer-reported data indicates children with ADHD are less socially preferred (Hoza et al., 2005) while teacher-reported data indicate children with ADHD are more likely to be rated as being less liked by their peers (Lahey et al., 1998). The emotional and behavioral dysregulation characteristic of children with ADHD has been suggested to increase the risk for being a victim of peer bullying (Unnever & Cornell, 2003). Parent, peer- and self-report data indicate children

with ADHD have fewer dyadic friendships compared to gender and age matched peers (Hoza et al., 2005; McConaughy et al., 2011), as children with ADHD endorse greater difficulties making and keeping friends compared to their peers (Lahey et al., 1998). Symptoms of hyperactivity/impulsivity have been associated with greater peer nominations of physical aggression and predictive of observed relational aggression (Diamantopoulou et al., 2007) suggesting that children with ADHD-HI and ADHD-C may experience more overt social impairments compared to children with ADHD-I.

The social skills deficits associated with ADHD also have negative implications for children's relationships outside of the school setting, as parents of children with ADHD are more likely to describe their child as having poorer relationships with family members and as being less involved in social activities (i.e., sports, clubs) (McConaughy et al., 2011). An investigation of patterns of social impairments by subtype of ADHD revealed children with ADHD-C demonstrate higher rates of emotional dysregulation and aggressive behavior while children with ADHD-I are characterized by a withdrawn and passive social style (Maedgen & Carlson, 2000).

Prevalence

Reported prevalence rates of ADHD in childhood range from 7.2% to 15.5% (Danielson et al., 2018a; Ramtekkar et al., 2010; Rowland et al., 2015; Thomas et al., 2015; Wolraich et al., 2014). Some degree of variability in these prevalence rates has been attributed to differences in the age of the children sampled, as there is evidence to suggest that prevalence rates of ADHD are highest in samples of preschool- and elementary-age children with decreasing prevalence

rates in older children (Willcutt, 2012). The high variability in prevalence rates of ADHD in childhood has also been attributed to a number of methodological limitations, such as calculating prevalence rates using samples that are not nationally representative and inconsistencies in calculating prevalence rates based on those individuals with a lifetime history of ADHD or limiting cases to only those individuals with a current diagnosis of ADHD. In an effort to overcome these methodological limitations, a team of researchers investigated data from the 2016 National Survey of Children's Health to examine prevalence rates of a nationally representative sample of children (ages 2 to 17) in the US and found an estimated prevalence rate of 9.4% (95% CI [8.8, 9.9]) among children with a lifetime history of ADHD and an estimated prevalence rate of 8.4% (95% CI [7.9, 8.9]) among children with a current diagnosis of ADHD (Danielson et al., 2018a).

Assessment of ADHD

Psychological assessment is a multipurpose procedure which can serve one or more functions depending on the goals of the assessment process. Such functions include: screening, diagnosing, treatment planning, treatment monitoring, and/or evaluating treatment outcomes (Mash & Hunsley, 2005). In clinical settings, assessments may primarily serve a diagnostic function, fulfilling the goal of identifying the an individual's psychopathology (Pelham et al., 2005). Assessments can also serve secondary functions of treatment planning and treatment evaluation which serve to identify the individual's specific treatment needs and assess the effectiveness of the selected treatment intervention(s), respectively (Pelham et al., 2005).

"Evidence-based assessment" (EBA) is a term often used to describe those assessment methods which have empirical evidence to support its clinical utility and diagnostic reliability and validity (Mash & Hunsley, 2005). Although there are established standards for the reliability and validity of individual assessment tools (Standards for Educational and Psychological Testing, 2014; Youngstrom et al., 2017), there are currently no established standards on how to curate an evidence-based psychological assessment battery (Mash & Hunsley, 2005). Consequently, there is no singular measure or set of measures which has been empirically demonstrated to be the gold-standard of assessment of ADHD in childhood (C. T. Barry et al., 2019). In the absence of such guidelines, mental health providers must use the principles of EBA to select those assessment tools which match the goal(s) of assessment, maximize accuracy, and remain feasible to implement.

Although DSM-5 symptom criteria are a cornerstone of diagnostic assessment, it has been argued that assessment batteries which are limited to measuring the presence of inattention and hyperactivity/impulsivity symptoms are insufficient to satisfy key assessment goals of treatment planning and treatment monitoring because symptoms of ADHD alone are not meaningful treatment targets (Pelham et al., 2005). Instead, assessment batteries should include tools which evaluate both behavioral symptoms *and* the way in which these dysfunctional behaviors lead to academic, emotional, and social impairment (Power et al., 2017). With a more comprehensive and accurate understanding of the child's psychopathology, treatment goals can be developed to address socially valid target behaviors. That is to say, target behaviors which, if improved upon, would be perceived as worthwhile and desirable

outcomes by society (Foster & Mash, 1999). For children with ADHD, socially valid target behaviors directly address psychosocial impairments (e.g., poor academic functioning, social functioning, family functioning) and/or deficits in adaptive skills, as these two domains of dysfunction are what lead children to be referred for evaluation and services and are the predictors of adverse outcomes across the lifespan (Pelham et al., 2005). Despite the high rates of psychological comorbidity among children with ADHD, it has been suggested that comorbid diagnoses themselves should not inform treatment planning and treatment evaluation (Pelham & Fabiano, 2001). Rather, the psychosocial impairments experienced by the child, subsequent to their idiosyncratic manifestations of one or more psychological disorders, should guide treatment planning and the development of socially valid treatment goals. To this end, the current study will implement EBA methods to evaluate (1) the suitability of the study intervention at pre-treatment and (2) the effectiveness of the study intervention at post-treatment by assessing children's: symptoms of ADHD and common comorbid internalizing and externalizing disorders, areas of psychosocial impairment, and deficits in adaptive skills.

Clinical practice guidelines for the diagnosis of ADHD in childhood (Felt et al., 2014; Wolraich et al., 2019) recommend the use of multimodal assessment tools (e.g., clinical interview, behavioral observations, and rating scales) and the use of multiple informants (e.g., parents, teachers). More specifically, it is recommended that assessment batteries for ADHD include: a clinical interview with the child's parents, behavioral observations of the child in one or more relevant settings, and behavior rating scales completed by both parents and teachers. In order to prepare a multimodal assessment battery that maximizes clinical utility and

efficiency, assessment tools should be carefully selected so that each instrument adds incremental validity to the overall assessment process (i.e., each tool provides unique data which is in service of meeting assessment goal) (Johnston & Murray, 2003). Considerations of which assessment modalities will be incorporated into an assessment battery, which discrete constructs of clinical interest will be assessed, and which informants will be asked to participate in the assessment process are all, in essence, considerations of incremental validity (Johnston & Murray, 2003). It is important to note that not all assessment modalities provide equal incremental validity. For instance, when parents and teachers are asked to complete rating scales regarding child behavior, there is no evidence for incremental validity or clinical utility in incorporating a structured clinical interview in an assessment battery for ADHD (Pelham et al., 2005). Furthermore, the incremental validity of a given assessment modality can vary depending on the goal of the assessment. For example, behavioral observations, while clinically useful for treatment monitoring purposes, are not considered to add incremental validity to diagnostic and treatment planning assessments (Pelham et al., 2005). In all, the theoretical promise that multimodal assessment batteries of greater rigor will lead to more robust assessment findings for externalizing disorders such as ADHD is one that has yet to be empirically proven (Johnston & Murray, 2003). It has been argued that combining simpler assessment tools (e.g., parent interview, parent- and teacher-reported behavior rating scales) is a an appropriate and efficient alternative to otherwise laborious and excessive assessment batteries because it honors the underlying principles of both multimodal assessments and

incremental validity (Johnston & Murray, 2003). Examples of such simpler assessment tools are described in further detail below.

Parent Interview. Clinical interviews are an essential component of assessments designed to screen and/or diagnose psychopathology in children (Barry et al., 2019). Unstructured clinical interviews, which are administered in a manner consistent with the clinical interviewer's level of expertise and the parents' presenting concerns, allow for a wealth of clinically relevant information to be discovered, including: onset of symptoms, precipitating events, settings in which symptoms lead to impairment, and previous interventions and outcomes (Barry et al., 2013; Mash & Hunsley, 2005). In turn, these data points facilitate differential diagnoses and strengthen treatment planning efforts (Barry et al., 2019).

A notable limitation to the use of unstructured clinical interviews is that they do not provide the interviewer with an accurate measure of how typical or atypical a child's symptom or impairment is in reference to same-aged peers (Barry et al., 2019). Additionally, unstructured clinical interviews are less consistent and reliable than structured clinical interviews that require the interviewer to systematically assess for a range DSM-5 disorders (Barry et al., 2019). However, structured clinical interviews are not essential for the diagnosis of ADHD, and the concurrent use of behavior rating scales can be used to indirectly address the limitation posed by an absence of age-based norms in unstructured clinical interviews (Barry et al., 2019).

In this study, an unstructured clinical interview was conducted at pre-treatment to inquire about children's primary problem behaviors, obtain parents' subjective report of the

psychosocial impairments experienced by their child, and to assess for children's skills deficits. To this end, the unstructured clinical interview was just one component of the assessment battery designed to satisfy the goals of screening and treatment planning.

Behavior Rating Scales. Behavior rating scales, particularly those that contain items rooted in DSM-5 symptom criteria and those with normed scoring guidelines, are considered to be a parsimonious and effective assessment method for ADHD (Achenbach et al., 2019; Barry et al., 2019; Pelham et al., 2005). Furthermore, rating scales that offer parallel forms (e.g., a parent report form and a teacher report form) allow the assessor to inquire about the same problem behaviors, with items phrased with behavioral descriptors that are most appropriate for the intended informant, from multiple informants. Cross-informant ratings are, on average, less correlated when informants have different relationships with the child in question and when informants are reporting on different settings in which the child has demonstrated behaviors (i.e., parents vs. teachers vs. clinicians; $r = 0.28$) (Achenbach et al., 1987). In comparison, cross-informant ratings are more highly correlated when informants have similar contextual relationships with the child (i.e., two parents in the same household or two or more teachers; $r = 0.60$) (Achenbach et al., 1987). To this end, discrepant reports are a rich source of unique, equally valid information. In fact, these noted discrepancies can help guide treatment planning to ensure appropriate goals and treatment interventions are applied to the appropriate settings (Achenbach et al., 2019). Accurate and comprehensive assessments of child psychopathology require data points from multiple informants as no single source of assessment data is sufficient to provide a full understanding of the child's functioning (Barry et

al., 2019). To this end, behavior rating scales are an invaluable assessment tool that enhances clinical utility and incremental validity of a given assessment battery.

Of note, behavioral rating scales can be broadband measures (i.e., measures which contain items designed to assess a number of psychological domains) or focused measures (i.e., measures that assess for a single psychological disorder) (Barry et al., 2019; Youngstrom, 2016). An example of a broadband measure is the Brief Problem Monitor (BPM; Achenbach et al., 2011). This brief, 19-item measure is an adaptation of the 113-item Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) and is validated to assess internalizing, externalizing, and attention problems in children (Piper et al., 2014). Of note, the attention subscale accurately identifies children with ADHD (Chen et al., 1994). Thus, the BPM is a broadband behavioral rating scale that can effectively discriminate children with significant internalizing, externalizing, and/or attention problems (Chen et al., 1994; Piper et al., 2014). In accordance with principles of EBA, focused measures are administered subsequent to broadband measures to reduce rates of false positive results produced by the broadband assessment (Youngstrom, 2016). One focused measure that has diagnostic specificity for ADHD is the Conners 3rd Edition – Short Form (Conners 3-(S); Conners, 2008). This measure offers parallel parent- and teacher-report forms and contains items specific to six content areas that are associated with ADHD: inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations. Both the BPM and the Conners 3-(S) rating scales are considered to be robust assessment tools as they provide normed *T* scores, accounting for empirical differences in the clinical presentation of ADHD by age, gender, and

culture (Achenbach et al., 2019). In this study, implementation of the BPM and the Conners 3-(S) rating scales at pre-treatment assessments satisfied the goals of screening and treatment planning, and the Conners-3-P(S) was administered again at post-treatment for the purposes of treatment evaluation.

Behavioral Observations. Although not considered to add incremental validity to diagnostic and treatment planning assessments (Lobitz & Johnson, 1976; Pelham et al., 2005; Sleator & Ullmann, 1981), behavioral observations have great clinical utility in regard to treatment monitoring (Pelham & Hoza, 1996). Specifically, observations allow for behavioral treatment targets identified in the treatment planning stage to be assessed over time, allowing for a qualitatively unique measure of a child's response to treatment (Willcutt et al., 1999). Serving the functions of monitoring and evaluating outcomes, behavioral observations are a key EBA tool (Youngstrom, 2017). This study implemented a standardized behavioral observation tool – Summer Treatment Program (STP) point system (Pelham et al., 2017; Synn et al., 2019) – to monitor the frequency of adaptive skills and positive and negative behaviors demonstrated by children over the course of the treatment, satisfying a key assessment goal of treatment monitoring.

Sex Disparities in ADHD Diagnosis

Across childhood and adolescence, a significantly greater proportion of males are diagnosed with ADHD compared to females. The ratio of males to females diagnosed with ADHD varies from 2:1 to 4:1 in community samples (Cantwell, 1996; Danielson et al., 2018; Pastor et al., 2015; Ramtekkar et al., 2010; Willcutt, 2012) and from 6:1 to 9:1 in clinical

samples (Cantwell, 1996; Gaub & Carlson, 1997). These findings have prompted further study of the factors responsible for this sex disparity with a secondary function of discerning the "true" degree to which there is a disparity in rates of diagnosis between males and females.

One commonly proposed explanation for the disparity of rates of diagnosis by sex is a referral bias, whereby symptoms of hyperactivity and impulsivity – which are more commonly reported as characteristic symptoms of ADHD in males and are more overt problem behaviors leading to distress for both the child and other individuals in their environment (e.g., peers, teachers, siblings, parents) – result in males to be referred for clinical evaluations of ADHD at higher rates than females (Abikoff et al., 1993; Biederman et al., 2005; Gaub & Carlson, 1997). In contrast, symptoms of inattention – which are more characteristic of females with ADHD and manifests as more subtle impairments in social and academic settings – are overlooked by teachers or parents (Gershon, 2002; Mowlem et al., 2019) or mis-conceptualized as symptoms of an internalizing disorder (L. E. Arnold et al., 2004; Cantwell, 1996). Consequently, the higher likelihood of referrals for clinical evaluation of ADHD in males and the underdiagnosis of ADHD in females are believed to culminate in disproportionate ratios as high as 9:1 in clinical ("referred") samples as compared to the ratios ranging from 2:1 to 4:1 in community ("non-referred") samples (Biederman et al., 2002; Gaub & Carlson, 1997; Wilcutt, 2012).

Racial and Ethnic Disparities in ADHD Diagnosis

In addition to the aforementioned sex disparities, researchers have also investigated whether there are disparities in the diagnosis of ADHD in childhood among different racial and ethnic groups. Some research suggests that non-Hispanic White children are more likely to

have a diagnosis of ADHD compared to both Hispanic and African American children (Coker et al., 2016; P. N. Pastor & Reuben, 2005; Rowland et al., 2002). (Most notably, a systematic review of the literature found that non-Hispanic White children were more likely to be diagnosed with ADHD while minority children were more likely to be diagnosed with a disruptive behavior disorder (Liang et al., 2016). These disparities may be reflective of both widespread underdiagnosis of psychopathology in Hispanic and African American children (Coker et al., 2016) and the limitations of current assessment methods in providing culturally sensitive, specific, reliable, and valid diagnostic information (Kirk, 2004; Lewis-Fernández & Díaz, 2002; Yeh et al., 2002).

In contrast, there is research to suggest that there are no significant racial and ethnic disparities in the diagnosis of ADHD in childhood. In a sample of elementary school children, researchers found no significant differences in the rates of ADHD diagnosis between non-Hispanic White and African American youth (Rowland et al., 2002). Additionally, in a sample of nationally representative children, there was no evidence of racial and ethnic disparities in the diagnosis of ADHD (Froehlich et al., 2007). It is important to note that these non-findings may be due to methodological limitations such as: (a) the omission of socioeconomic variables as covariates in the statistical analyses (Coker et al., 2016 and (b) obtaining diagnostic data solely from parents, a notable limitation from an EBA perspective as multiple informants with different relationships with the child (e.g., parent and teacher) are necessary to obtaining a comprehensive record of the child's presenting symptoms and level of impairment (Barry et al., 2019; Mash & Hunsley, 2005). In light of these mixed findings, there is a need for further

research investigating the degree to which racial and ethnic disparities in ADHD diagnosis persist and the factors contributing to this disparity.

Course of ADHD Across the Lifespan

The existing literature conservatively estimates that at least two-thirds of children diagnosed with ADHD will continue to demonstrate persistent symptoms of ADHD in adulthood (Biederman, 1996; Biederman et al., 2011; Mick et al., 2011; Sibley et al., 2017). Although there is a small subset of children that may experience remission of symptoms at some point in their life, there is consistent evidence to suggest that ADHD is a chronic neurodevelopmental disorder for a majority of diagnosed individuals. Across the lifespan, predictors of persistence (i.e., of continuing to meet diagnostic criteria, per the DSM, for ADHD) include: a greater number of ADHD symptoms in childhood, a family history of ADHD, parental history of psychopathology, marital conflict, and a comorbid disruptive behavior, anxiety, and/or mood disorder (Biederman et al., 1996, 2011; Roy et al., 2016). However, in childhood, there are other unique predictors of persistence, such as: parents' use of inconsistent discipline strategies, a history of early family adversity (e.g., medical illness of a family member, parental divorce or separation, financial stressors, legal stressors), and low family income status (Lahey et al., 2016; Sasser et al., 2016).

Negative Sequelae of Persistent ADHD

Given the likelihood of persistent ADHD across the lifespan, a life transition model for ADHD – mirroring models that have been developed for other chronic physical and mental health conditions – was developed to promote continuity of care across different developmental

stages and to decrease the risks for adverse sequelae of ADHD (Turgay et al., 2012). The life transition model specific to ADHD presupposes two "sensitive periods" in youth development which carry implications for long-term outcomes: (1) the transition from childhood to adolescence and (2) the transition from adolescence to young adulthood. In each of these transitional periods, the individual is faced with increased demands from the environment (i.e., academic, social and/or financial). In the absence of receiving the requisite support to meet these increasing demands, the individual experiences functional impairment and a higher likelihood of long-term adverse outcomes.

Evidence for these sensitive periods have emerged from longitudinal studies examining adverse sequelae of ADHD across childhood, adolescence, young adulthood, and beyond. Specifically, the presence of ADHD in childhood has been shown to place youth at heightened risk for a variety of poorer academic outcomes in adolescence, including: a greater number of suspensions, grade retention, receiving lower grades, expulsions and dropout from high school, and failing to enroll in college (Barkley et al., 2006; Owens & Hinshaw, 2016). Additionally, individuals with ADHD in childhood have been shown to be at increased risk for arrests, unintentional injuries, and risky motor vehicle behaviors in adolescence (Lahey et al., 2016). Studies have also demonstrated that ADHD in childhood and adolescence is predictive of: impairments in peer relationships (Barkley et al., 2006), poorer occupational attainment and impaired occupational functioning (Barkley et al., 2006), greater need for public assistance (Hechtman et al., 2016), risky sexual behavior (Nigg, 2013), risky driving (Lahey et al., 2016), substance use disorders (Nigg, 2013), crime and arrests (Lahey et al., 2016), unintentional

injuries (Lahey et al., 2016; Nigg, 2013), elevated health risks including obesity (Cortese et al., 2013; Nigg, 2013), eating pathology in females (Mikami et al., 2008), and cardiovascular disease (Nigg, 2013), greater use of health care systems (Nigg, 2013), and greater number of years of disability and ill-health (Barkley & Fischer, 2019) in adulthood and early mortality (Dalsgaard et al., 2015; Faraone, 2015). Of note, the diagnosis of ADHD is not predictive of a reduced lifespan, in and of itself. Rather, this neurobehavioral disorder engenders persistent inattention and impulsivity in an individual's life which is a risk factor for antisocial behaviors, substance use, and other risky behaviors. In turn, these are risk factors for violence, crime, accidents, and poor health habits that result in higher likelihood of premature death (Barkley & Fischer, 2019; Dalsgaard, 2015; Faraone, 2015).

Given the multiplicative nature of ADHD-related adversity, there has been a growing impetus to prevent persistence of ADHD symptomatology and impairment in adulthood (Turgay et al., 2012) and to promote remission rates in childhood and adolescence (Steele et al., 2006). Remission of ADHD is associated with reduced functional impairments in academic, social, and occupational domains (Hechtman et al., 2016). Remission of ADHD symptoms is also associated with decreased symptomatology of anxiety and depression and decreased rates of illicit drug use compared to individuals with persistent ADHD (Young & Gudjonsson, 2008).

Common Comorbidities

Nearly two in three children with ADHD have a comorbid mental, emotional, or behavioral disorder (Cuffe et al., 2015). Common comorbid conditions among children with ADHD include: oppositional defiant disorder (ODD; Cuffe et al., 2015, Elia et al., 2008; Tung et

al., 2016), conduct disorder (CD; Cuffe et al., 2015, Tung et al., 2016), anxiety (Cuffe et al., 2015, Elia et al., 2008; Tung et al., 2016), depression (Cuffe et al., 2015; Tung et al., 2016), and dysthymia (Elia et al., 2008). The high rates of psychiatric comorbidities among children with ADHD result in this pediatric population being of greater public health concern: in addition to the functional impairments stemming directly from their neurodevelopmental disorder, it is less often the exception and more often the rule that these youth will experience intensified adversity due to the nature of their comorbid condition(s). Recent evidence suggests children with ADHD and a comorbid anxiety or mood disorder have a significantly increased risk for below average academic performance (OR = 59.9; Cuffe et al., 2015). In contrast, youth with ADHD and a comorbid disruptive behavior disorder, such as ODD or CD, have higher odds of trouble with police, school expulsion, suspension, and/or in-school suspension (OR= 14.1; Cuffe et al., 2015), and the presence of conduct problem behaviors in childhood are predictive of poorer overall functioning and greater internalizing problems in young adulthood (Owens & Hinshaw, 2016).

Development of Comorbid ADHD and Depression. Functional impairments subsequent to ADHD – such as poor academic functioning (Cuffe et al., 2015; Loe & Feldman, 2007), strained relationships (Biederman et al., 1998; Daviss, 2008), and negative parent-child interactions (Daviss, 2008) – are believed to engender significant risk for developing a comorbid depressive disorder. Specifically, the pervasiveness of ADHD-related functional impairments is proposed to foster negative evaluations about self and circumstances, that culminate in a state of persistent demoralization (Biederman et al., 1998). In the absence of

timely intervention, the heightened distress and despair can give rise to a depressive disorder. In support of this sequential development of comorbidity, the presence of ADHD in adolescence – independent of gender, other psychological comorbidities, or level of impairment in academic and social settings – has been associated with a greater risk of developing major depressive disorder (MDD) in adulthood (Meinzer et al., 2013). Additionally, persistent symptoms of ADHD, that in turn lead to persistent distress and functional impairment, is associated with an increased life-time risk for depression (Riglin et al., 2020).

Development of Comorbid ADHD and Oppositional Defiant Disorder. In an effort to explain the high rates of comorbidity between ADHD and ODD, two models have recently been proposed and empirically examined. The developmental precursors model (E. A. Harvey et al., 2016) posits that symptoms of ADHD lead to specific child behaviors that invariably transact on the family system, and subsequent disruptions in family dynamics lead to the development of argumentative and defiant symptoms characteristic of ODD. Alternatively, the correlated risk factors model (Harvey et al., 2016) posits that a family history of ODD or CD has unique predictive power on the development of anger and irritability symptoms characteristic of ODD or CD in a child with ADHD. Both models have demonstrated predictive validity of the development of comorbid ODD, suggesting that there are multiple pathways by which youth with ADHD can develop ODD later in life (Harvey et al., 2016). Of note, there is also preliminary evidence to suggest that specific constellations of ADHD symptoms may increase a child's risk for ODD. Specifically, children with ADHD-HI and ADHD-C, as compared to children with

ADHD-I, have been shown to be more likely to have a comorbid diagnosis of ODD (Elia et al., 2008).

Summary of ADHD in Childhood

ADHD is a chronic neurodevelopmental disorder (APA, 2013) affecting approximately 8.4% of children in the US (Danielson et al., 2018a). In childhood, ADHD is the most prevalent reason for referral to a pediatric or mental health professional (Barkley, 1990). Children with ADHD commonly have a co-occurring mental, emotional, or behavioral disorder (Danielson et al., 2018a) and experience pervasive distress and impairment in both academic and social settings (Barbaressi et al., 2007a; Cervantes et al., 2013; Dimatopoulous et al., 2007, Hoza et al., 2005; LaFever et al., 2002; Lahey et al., 1998). In the absence of timely, evidence-based intervention, children with ADHD are more likely to experience peer rejection and poorer academic attainment and attendance (Barbaressi et al., 2007b; Biederman et al., 1996). Long-term consequences of inadequate treatment of ADHD in childhood include a higher likelihood of engaging in high risk behaviors in adolescence (Lahey et al., 2016; Nigg, 2012) and of unemployment, poorer physical health outcomes, and early mortality in adulthood (Barkley et al., 2006; Barkley & Fischer, 2019; Daslgaard, 2015; Farone, 2015). Similar to other chronic physical and mental health disorders, childhood and adolescence have been identified as two critical periods for the assessment and treatment of ADHD (Turgay et al., 2012). It has been argued that treatment in childhood presents the greatest chances of robust long-term outcomes (Steele et al., 2006; Turgay et al., 2012). To this end, it is of significant clinical importance to implement evidence-based assessments and treatments in childhood to

effectively manage ADHD and mitigate the negative sequelae of this chronic disorder across the lifespan.

Empirically Supported Treatments for Children with ADHD

The following sections describe empirically supported, or evidence-based, interventions for the treatment of ADHD in childhood. In the context of treatment, "evidence-based" is used to describe those interventions with scientific evidence to support its safety and efficacy in promoting psychological health (APA, 2006). The current evidence-base for the treatment of ADHD in childhood consists of three categories of treatment approaches: (1) psychosocial treatments, (2) psychopharmacological treatments, and (3) combined treatments, which involve the concurrent use of psychosocial and psychopharmacological treatments (Brown et al., 2008).

Psychosocial Treatment Interventions

Psychosocial treatment approaches to ADHD in childhood are those that act directly on the child or the child's direct environment (e.g., parents, home, teachers, school system) to promote desired changes in the child's symptomatology and functioning. Among the range of psychosocial interventions that have been investigated for use with children with ADHD, only a limited number of interventions have demonstrated: (1) sufficient rigor of methodology and (2) efficacy in at least two independent research settings by two independent investigatory teams (Evans et al., 2018). Interventions that meet criteria for this highest level of evidence are deemed "well-established" interventions (Evans et al., 2018; Southam-Gerow & Prinstein, 2014). Such interventions for ADHD, include: behavioral parent training (BPT), behavioral classroom management (BCM), and intensive behavioral treatment which is a multimodal

behavioral intervention incorporating both BPT and BCM interventions (Chronis et al., 2006; Evans et al., 2018; Fabiano et al., 2014; Pelham & Fabiano, 2008). Given the specific focus on well-established interventions for ADHD in childhood, social skills training and cognitive behavioral therapy (CBT) interventions will consequently be omitted from the subsequent review.

Common Principles. Common to each of the well-established psychosocial treatment interventions for childhood ADHD are two noteworthy principles that account for effectiveness in helping the child learn new behaviors: operant conditioning (Skinner, 1963) and social learning theory (Bandura, 1977). These principles are described in further detail below.

Operant Conditioning. Operant conditioning emerged from Thorndike's Law of effect (1911), which stated:

> Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it reoccurs, they will be more likely to reoccur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with the situation weakened, so that, when it reoccurs, they will be less likely to occur. The greater the satisfaction of discomfort, the greater the strengthening or weakening of the bond. (p. 244)

Thus, to a large extent, behavior modification and behavior management protocols rely on control of consequences to increase or decrease the probability of a behavior occurring in the future. In the case of childhood psychopathology, operant contingencies are carefully

examined, monitored, and modified in order to alleviate future risk for continued psychopathology and its associated distress, functional impairment, and cost to society at large. Several studies have investigated the effectiveness of varying reinforcement schedules for a range of clinical populations, leading to a strong evidence base for the use of operant conditioning principles to promote positive behavior change (Farmer & Chapman, 2016; Luman et al., 2005; Staddon & Cerutti, 2003). Although the principles underlying clinically recommended behavior management strategies are universal and can be implemented in a tailored fashion to meet the needs of children with an array of presenting problem behaviors, it is important to note that some reinforcement schedules have been shown to be more effective among certain clinical populations. For example, a salient difference between children with ADHD and their non-ADHD peers has been observed: children with ADHD are more likely to prefer immediate reinforcement while delayed reinforcement remains an effective behavior management strategy for non-ADHD peers (Luman et al., 2005; Sagvolden et al., 1998; Sonuga-Barke, 2003).

Social Learning Theory. Social learning theory posits that new behaviors are learned by either engaging in the behavior (i.e., direct experience) or witnessing another individual engage in the behavior (i.e., modeling) (Bandura, 1971). Although the principles of social learning do not demand the concurrent application of reward and punishment strategies in order for learning to occur, the principles of operant conditioning and social learning augment each other to optimize behavioral learning. For instance, if a child was being taught how to sit appropriately in a classroom, the direct experience of that target behavior would result in a

basic understanding of what body posture the teacher desires of her students. However, this learning is amplified when the child is rewarded with a desirable token, item, or privilege for engaging in the target behavior as the timely consequence provides immediate feedback to the child. As social learning theory further postulates, children can also learn from others' experiences (Bandura, 1971). If a child observes peers receiving rewards for appropriate sitting behaviors, the child experiences vicarious reinforcement and learns which behaviors are desired in the academic setting. This, in turn, increases the likelihood of the child demonstrating appropriate sitting behavior in the future. Conversely, behaviors which are observed to result in punishing consequences can discourage the child from emulating a behavior.

Behavioral Parent Training. The early childhood coercion model serves to explain how difficult child behaviors and ineffective parenting strategies reciprocally influence each other and can lead to coercive parent-child interactions which inadvertently maintain or exacerbate child problem behaviors (Scaramella & Leve, 2004). BPT treatment protocols serve to remediate these parent-child transactions by first rehabilitating parent-child relationships and then instructing and coaching parents on how to promote adaptive child behaviors by modifying the child's environment (*antecedent strategies*) and/or strategically implementing rewards and punishments (*consequence strategies*) (Chronis et al., 2004). BPT sessions are commonly completed with the parent(s) alone. However, some sessions may include the child as this provides an opportunity for the parent to receive live-coaching from the clinician on the use of effective parenting strategies, thereby enhancing parents' understanding and mastery of skills (Kazdin, 2009). BPT is also commonly delivered in a group format as a means of enhancing

parents' sense of support and accessibility to families in need (Kazdin, 1997; Pelham & Hoza, 1996).

Effectiveness of BPT. Compared to psychosocial interventions where parents receive unstructured support based on their primary areas of concern, research suggests BPT is uniquely potent and effective in ameliorating symptoms of ADHD in childhood due to the explicit training in parenting strategies (e.g., active ignoring, use of praise, effective commands and reminders) it provides (Sonuga-Barke et al., 2001). Additionally, there is evidence to suggest that the primary focus of fostering a more positive parent-child relationship serves to make BPT a potent remedy for children's ADHD-related maladaptive behaviors and functional impairment (Deault, 2010). In samples of preschool children, the positive outcomes from BPT were observed to last for up to 15 weeks after the conclusion of the treatment intervention (Sonuga-Barke et al., 2001). The positive treatment effects of BPT can be extended upwards of 9 months when it is part of a comprehensive psychosocial treatment package (L. E. Arnold et al., 2004). To this end, current clinical practice guidelines indicate BPT as a first-line intervention for preschool and elementary age children with ADHD (Evans et al., 2018; Sonuga-Barke et al., 2013; Wolraich et al., 2019) with additional supports (e.g., BCM, psychostimulants) added on adjunctively, as needed.

Predictors and Moderators of BPT Treatment Outcomes. Despite the robust evidence in support of BPT as a primary intervention for children with ADHD, more recent research efforts have sought to investigate what child and parent characteristics, if any, may predict or moderate BPT treatment outcomes. There is evidence to suggest that certain patterns of child

psychological comorbidity may moderate treatment outcomes. For example children with two or more psychological comorbidities, as compared to children with one or no comorbidities, demonstrate greater rates of treatment success when prescribed a behavioral intervention for ADHD (van den Hoofdakker et al., 2010). There is also evidence to suggest that children with comorbid ADHD and anxiety ("ADHD+ANX") demonstrate greater parent-reported treatment success in response to behavioral intervention compared to children with ADHD only (MTA Cooperative Group, 1999).

Although parents of different ethnicities have been shown to demonstrate different parenting styles at baseline, parental ethnicity has not been found to be a predictor or moderator of treatment outcome (Jones et al., 2010). As such, there is evidence that BPT is a culturally sensitive treatment intervention that is equally efficacious for children and parents of different ethnicities (Jones et al., 2010). In contrast, there is evidence to suggest that parenting self-efficacy (i.e., a parent's positive cognitions about their abilities as a parent to meet their child's needs) is a moderator of positive treatment outcomes for both mothers (van den Hoofdakker et al., 2010) and fathers (van den Hoofdakker et al., 2014) (who complete a BPT treatment protocol. More specifically, greater parenting self-efficacy at baseline is associated with greater chances of response to a BPT intervention. Further research is needed to discern if, and to what extent, levels of parental self-efficacy may increase during the course of a BPT intervention.

Another investigated moderator of BPT treatment outcomes is parental psychopathology at baseline. While some studies conclude there is no moderating effect of parental

psychopathology (e.g., ADHD, depression) on treatment outcomes (Van den Hoofdakker et al., 2010; Van den Hoofdakker et al., 2014), other studies have demonstrated evidence of a moderating effect (Chronis-Tuscano et al., 2011; Gerdes et al., 2007; van den Hoofdakker et al., 2014). These discrepant findings may be due to the presence of varying intensities of parental psychopathology, whereby only certain levels of functional impairment lead to diminished parental self-efficacy or significant interference in consistently implementing effective parenting strategies prescribed by the BPT protocol. Additionally, by virtue of being two discrete disorders, parental ADHD and parental depression likely mediate BPT treatment outcomes in unique ways, leading to the discrepant findings regarding the moderating effects of parental psychopathology.

Among parents with ADHD, a clinically significant number of inattentive and impulsive symptoms has been associated with permissive parenting (E. Harvey et al., 2003), a parenting style characterized by high levels of warmth and low levels of limit setting (Baumrind, 1967, 1971). Additionally, parental ADHD is associated with less positive parent-child relationships, as evidenced by more negative and critical statements during child-led play and inconsistent use of discipline strategies (Chronis-Tuscano et al., 2008). Although BPT serves to replace such ineffective parenting behaviors with more adaptive techniques, parents with ADHD have been shown to demonstrate higher rates of adverse parenting behaviors in spite of direct instruction and coaching on more effective parenting strategies (Chronis-Tuscano et al., 2011). Thus, the neurological and functional impairments associated with ADHD commonly lead to predictable parent behaviors that interfere with optimal treatment outcomes.

Depressed parents face distinct, yet equally notable, barriers to engagement with BPT intervention and adherence to treatment recommendations. The behavioral and cognitive sequelae of depression are believed to diminish depressed parents' ability to optimally engage in treatment, as evidenced by lower rates of adherence to prescribed therapeutic strategies (Owens et al., 2003). As such, it has been suggested that BPT treatments include parental stress and negative parental cognitions – both as they relate to their child– as targets of treatment (Chronis et al., 2004). Such flexible applications of the BPT intervention have been shown to directly and indirectly ameliorate levels of parenting stress as parents develop proficiency and mastery of more effective behavior management techniques and child behavior improves (Gerdes et al., 2012; Owens et al., 2003).

Behavioral Classroom Management. Broadly defined, behavioral classroom management (BCM) is the systematic application of reinforcement, punishment, and extinction operations to manage child behavior in the classroom setting (Landrum & Kauffman, 2006). The daily report card (DRC) is one specific BCM intervention and is commonly recognized as a best practice for youth with ADHD (Fabiano et al., 2010; Volpe & Fabiano, 2013). The DRC is characterized by individualized target behaviors which teachers monitor and assess during the school day. The child's level of success on the DRC – defined as the percentage of target behaviors achieved in one day – is then translated into either a home-based reward or punishment to be delivered by the parent.

The DRC intervention is an effective way to enhance a child's academic performance in the absence of any other form of psychosocial or pharmacological support (Kelley & McCain,

1995). More specifically, the DRC intervention reduces rates of child disruptive behavior in the classroom and promotes greater accuracy on completed assignments (Fabiano et al., 2010). Students who received the DRC intervention as part of an enhanced special education intervention are more likely to reach their individualized education plan (IEP) goals (Fabiano et al., 2010). However, in samples of youth with ADHD, the DRC as a unimodal intervention approach is not sufficient to reduce the number of ADHD symptoms or levels of impairment experienced by a child (Fabiano et al., 2010). Thus, complementary treatment strategies are often concurrently delivered to children with ADHD to address all areas of impairment and distress.

Intensive Behavioral Treatment. Combining behavior management strategies utilized in both BPT and BCM interventions, the summer treatment program (STP) is an intensive, multimodal behavioral treatment and a discrete evidence-based treatment approach. The STP was designed to promote adaptive functioning among youth with ADHD by: shaping more positive interactions with adults, fostering more positive peer relationships, preventing against summer academic regression, improving classroom behaviors and academic achievement, and enhancing children's sports skills training (Pelham & Hoza, 1996). To meet these treatment goals, the STP implements: (1) an intricate reward and response-cost system, (2) a time out protocol, and (3) the DRC intervention.

The STP reward and response-cost system targets a range of positive and negative behaviors, such that adaptive behaviors can be shaped and reinforced and maladaptive behaviors specific and common to children with ADHD, ODD, and CD can be punished. To this

end, youth in the STP are awarded points for positive behaviors (i.e., prosocial and on-task behaviors) and are docked points for negative behaviors (i.e., disrespectful, dangerous, and destructive behaviors). Of note, behavior management interventions that rely solely on reinforcement of adaptive behaviors are insufficient in promoting significant improvements in STP target behaviors (Pelham et al., 1993). In contrast, response-cost systems – which allow for the immediate delivery of positive and negative consequences, as needed – facilitate greater rates of behavior change.

During the STP, youth who engage in acts of aggression and/or destruction are met with two punishing consequences: a loss of points and an assigned time point. In this way, more egregious problem behaviors are met with sufficiently punishing consequences. The duration of the time out procedure is age-dependent and can be reduced in half if the child demonstrates compliance (i.e., the child immediately goes to the designated time out space without additional prompting). Of note, youth are intentionally left unaware of the time limits associated with the STP time out protocol. The conclusion of each time out is contingent on the child's behavior: after the duration of the time out has elapsed and the child has demonstrated at least 5 seconds of a calm and quiet body, the child is informed that their time out is over. Over time, this time out protocol shapes children to quickly comply with their assigned time out. Furthermore, the consistency and predictability of the time out protocol facilitates children's emotional and behavioral re-regulation during the time out.

In addition to the reward and response-cost system and time out protocol, the STP implements an additional layer of behavior management in the form of the DRC. This lends a

more individualized approach to contingency management for children participating in the STP. In this way, target behaviors spanning group settings, dyadic peer interactions, and classroom settings, are carefully selected to meet the more specific treatment goals of each individual child.

Effectiveness and Acceptability of the STP. Pelham and Hoza (1996) conducted preliminary analyses of the efficacy of the STP using a sample of 258 male children with ADHD who attended the program between the years 1987 and 1992. Primary outcomes were assessed using counselor-, parent-, and teacher-reported rating scales of ADHD and related behaviors at post-treatment. Counselor-reports indicated 91% of child participants were at least "somewhat improved", and parent-ratings indicated significant treatment effects in reducing both ADHD symptoms and symptoms of ODD and CD. Per anonymous parent reports, nearly 100% of respondents reported the STP to be at least "somewhat beneficial" for both their child and themselves. Additionally, approximately 95% of child participants were reported to have found the program enjoyable. Since this preliminary analysis, the robust acceptability and effectiveness of the STP intervention has been replicated by numerous investigative teams (Chronis et al., 2004; Fabiano et al., 2014; Pelham et al., 2000, 2005, 2014, 2016). Many of these studies compared the efficacy of the behavioral intervention (STP) with that of combined intervention (STP + concurrent stimulant pharmacotherapy). Study findings indicate treatment effects are comparable in relation to objective measures of behavior such that children do not demonstrate statistically significantly different outcomes based on treatment group (Chronis et al., 2004; Pelham et al., 2000). Treatment effects are also comparable in relation to DRC

outcomes: on average, children demonstrate a 10% increase in DRC scores from week 1 (75%) to week 8 (85%) (Pelham et al., 2000).

Several studies have also examined the incremental effects of different STP treatment components such as the academic learning center (Fabiano et al., 2007) and the time-out protocol (Fabiano et al., 2004). It is also important to note that response-cost systems, such as the one utilized in the STP protocol, have a distinct advantage over psychostimulant interventions in improving social behaviors, on-task behaviors, and completion and accuracy rates on assignments in the academic setting (Rapport et al., 1982). Children who participate in the STP have been shown to demonstrate significant improvements in social, recreational, and academic functioning, regardless of their gender, age, psychiatric comorbidities, and concurrent psychopharmacology treatment status (Coles et al., 2005). Of note, improvements in recreational functioning is characterized by: greater competencies in rules-based sports activities, improvements in sports performance while following activity rules, and enhanced sportsmanship (O'Connor et al., 2014). To this end, there is evidence to support the use of daily sports skills activities in the STP intervention as an effective modality of learning and rehearsing adaptive behaviors and decrease levels of overall functional impairment.

A recent investigation of the STP compared the traditional, high-intensity STP (e.g., 8-week program, providing 45 direct hours of intervention each week) to low-intensity summer interventions (e.g., 8-week program, providing 1.5 hours of group-based organization skills training to children once weekly), both of which also provided 1.5 hours of group-based parent training once weekly (Sibley et al., 2018). Objective measures of child behavior outcomes

indicated participants of the high-intensity intervention experienced significantly greater reductions in parent-reported ADHD symptoms. Additionally, youth attendance rates were higher in the high-intensity group as compared to that of the low-intensity group (approximately 80% and 45%, respectively). In all, the standard STP – which provides intensive behavioral treatment in a proxy environment for peer, social, academic, and home settings (Wells et al., 2000) – is a potent treatment intervention for a range of functional impairments resulting from chronically inattentive, overactive, and/or impulsive behaviors (Fabiano et al., 2014). However, the STP is currently only offered at one of fifteen sites across the US, and the 8-week treatment protocol often involves substantial costs, monetary and otherwise, for participating families.

Adaptations to Standard STP Model. In efforts to expand the feasibility and accessibility of the STP, adaptations to the standard STP protocol have been examined. A team of researchers in Japan conducted preliminary effectiveness studies of 2- and 3-week long STP programs, adapted to meet the needs of a shortened summer holiday. These adapted STPs were similar in structure to the standard STP with the exception of total daily classroom time: in the 3-week treatment protocol, children spent a total of 1 hour per day in a classroom environment, whereas the 2-week treatment protocol called for 1.5 hours per day in a classroom environment (Yamashita et al., 2010, 2011). Analyses concluded that the adapted STPs remained a sufficiently intensive treatment as evidenced by significant parent-reported reductions in symptoms of ADHD and ODD (Yamashita et al., 2010, 2011). Additionally, the shorter duration of the STP intervention, which required a less time-intensive commitment from

participating families, was believed to contribute to the high rate of retention (i.e., 100% adherence) and no dropouts from treatment (Yamashita et al., 2010, 2011). However, the 3-week STP did not produce significant improvements in parent-reported measures of peer relationships and prosocial behaviors (Yamashita et al., 2010) and additional research is needed to investigate the long-term outcomes associated with this condensed STP model. More recently, a team at Judge Baker Children's Center in Boston, MA developed a 6-week adaptation of the STP protocol. Preliminary data has demonstrated the 6-week STP to be feasible, accessible, and effective in reducing the severity of children's internalizing, externalizing, and attention problems (Tannenbaum et al., 2019). The current study seeks to contribute to the existing treatment outcome literature by expanding upon what is known about the clinical utility of the 6-week adaptation of the STP model delivered in a community setting.

Psychopharmacological Treatment Interventions

The US Food and Drug Administration (FDA) has approved two classes of medications for the treatment of ADHD in children ages 6 and older: psychostimulants ("stimulants") and non-psychostimulants ("non-stimulants") (US FDA, 2016). The most recent clinical practice guidelines for the treatment of ADHD in childhood and adolescence set forth by the American Academy of Pediatrics (AAP) explicitly states that FDA approved medications for ADHD should only be used as a first-line intervention for adolescent populations (ages 12 to 18) (Wolraich et al., 2019). Among preschool (ages 4 to 6) populations, the use of the stimulant intervention methylphenidate is recommended only when evidence-based behavioral interventions (e.g., BPT, BCM) are not sufficient in providing significant improvement; and among elementary and

middle school age children (ages 6 to 12), any one of the FDA-approved classes of medications for ADHD is only recommended for use in conjunction with evidence-based behavioral interventions (Wolraich et al., 2019). Compared to non-stimulant mediations, stimulants have been shown to be more efficacious and safe for the treatment of ADHD in childhood and adolescence (Felt et al., 2014). As such, the FDA recommends the use of stimulants as the first-line pharmacological intervention, with non-stimulants prescribed only as alternative pharmacological intervention for those children who are unable to tolerate stimulants (Felt et al., 2014; US FDA, 2016). To this end, the scope of the subsequent review will be limited to the evidence-base for stimulant medications for the treatment of ADHD in childhood.

Stimulant medications are prepared in one of three formulations: methylphenidate (MPH), dextroamphetamine (DEX), and mixed amphetamine salts (i.e., dextroamphetamine/amphetamine) (Felt et al., 2014), and stimulants are characterized by their effect on dopaminergic pathways in the central nervous system (Arnsten, 2006). Stimulant medications are among the most effective medicines in all of modern medicine, as upwards of 70% of youth with ADHD demonstrate positive response to this pharmacological intervention (Spencer et al., 1996). Stimulants have demonstrated a range of positive effects for youth with ADHD, namely decreasing risk for grade retention (Barbaressi et al., 2007b; Biederman et al., 2008) and increasing scores on standardized math and reading assessments (Scheffler et al., 2009). While the enhanced scores on standardized assessments (e.g., 2.9 points greater of a score on mathematics; 5.4 points greater of a score on reading) subsequent to stimulant intervention may appear to be marginal advancements in academic achievement, these higher

scores are equivalent to those that would have been achieved with approximately 34 to 52 days of additional schooling (Scheffler et al., 2009). As such, stimulant intervention can promote short-term academic success and facilitate continued patterns of academic success in the future. There is also evidence to suggest that the neural mechanism of action of stimulant interventions is one that may allow for situational dependency: for example, children treated with stimulants can demonstrate reduced inappropriate behavior in the classroom and appropriately low levels of motor activity during seatwork periods while still generating appropriately high levels of motor activity during a recess period within the same school day (Swanson et al., 2002). Additionally, children treated with stimulant pharmacotherapy have been shown to experience improvements in social functioning, as evidenced by reduced negative interpersonal interactions with peers, parents, and teachers (Abikoff et al., 2004). To this end, there is robust evidence for the use stimulant interventions in facilitating both clinically significant and subjectively valuable improvements in youth functioning.

In light of the high rate of psychological comorbidity among youth with ADHD, it is also important to note that stimulant interventions have been shown to decrease the risk of onset of comorbid disorders. A longitudinal study provided evidence that the use of stimulants during childhood and adolescence, between the ages of 6 and 18, can decrease the likelihood of experiencing an onset of anxiety disorders or disruptive behavior disorders up to 10 years later (Biederman et al., 2009). Recent research has also demonstrated the protective nature of psychostimulants in regards to the onset of a depressive episodes (Biederman et al., 2009; Chang et al., 2016). It is hypothesized that by ameliorating those ADHD-related symptoms

which result in persistent distress and impairment, there is a decreased likelihood of developing clinically significant levels of worry in and/or avoidance of evaluative situations, irritability, sadness, hopelessness, worthlessness, and/or feelings of guilt, which, in turn, can engender secondary pathology.

Moderators of Stimulant Treatment Outcomes. It is important to note that stimulants are not equally effective for all children with ADHD. For example, greater severity of child ADHD symptoms at baseline has been associated with lower rates of successful response to psychostimulant treatment (Owens et al., 2003). Additionally, parental depression at baseline has been associated with lower rates of child response to stimulant intervention (Owens et al., 2003). In the case of severe child symptoms of ADHD and parental depression at baseline, below average child IQ has also been identified as a moderator of poorer treatment response to stimulant intervention (Owens et al., 2003). There has also been research to suggest that children with different psychological comorbidities will respond differently to stimulant intervention for the treatment of ADHD. More specifically, children with ADHD and comorbid ODD or CD ("ADHD+ODD/CD") demonstrate greater treatment response rates to stimulant interventions than children with comorbid ADHD, anxiety, and ODD or CD ("ADHD+ANX+ODD/CD") (Jensen et al., 2001). To this end, the degree of child ADHD severity, pattern of psychological comorbidity, and presence of parental depression are important clinical considerations when establishing a treatment plan to address ADHD in childhood.

Concerns of Diversion and Abuse. Despite strong evidence for the use of stimulants in ameliorating ADHD-related symptomatology and functional impairment, a common concern

regarding stimulant interventions is the risk for substance abuse, addiction, and substance use disorder across the lifespan. Studies have shown that when stimulant medications are prescribed by licensed professionals, therapeutic doses of medication are maintained, and administration of medication is carefully monitored, stimulant medications are associated with increasingly protective effects against substance abuse and substance use disorder over time (Chang et al., 2014). Additionally, stimulants, which target the reward circuitry in the brain, directly reduce symptoms of impulsivity, which otherwise confers risk for substance use (Faraone & Wilens, 2007). To this end, effective stimulant intervention – where dosing is carefully adjusted by a licensed physician to match changes in the child's metabolism, growth, and severity of symptoms – mitigates risk for illicit substance use (e.g., youth seeking to self-medicate as a means of alleviating distressing symptoms of impaired mood or sleep subsequent to their psychological disorder) (Wilens, 2008). Of note, there is also research to suggest that parents who seek psychiatric support out of concern for their child are more likely to provide adequate supervision of the controlled substance, thereby reducing risks for misuse of the prescribed medication (Faraone & Wilens, 2007).

Combined Treatment Interventions

To date, investigations of combined intervention versus medication or behavioral therapy alone have been largely limited to primary and secondary analyses of the seminal, NIMH-funded Multimodal Treatment Study of Children with ADHD (the MTA study) (Arnold et al., 1997). The MTA study (n = 576 children) was conducted across six sites over a period of 14 months. Child participants, ages 7 to 9, were randomized to one of four treatment conditions:

(1) medication alone (MPH as the first-line intervention; and open titration with other medications, as needed, if an effective dose of MPH was not found within 1 month), (2) psychosocial treatment alone (multicomponent intervention including: parent training, teacher consultation, half-time behavior specialist intervention in the classroom, and child participation in an 8-week STP), (3) combination treatment (integration of both the medication and psychosocial treatment strategies), and (4) community comparison (child receives treatment from a community provider of their own choosing). An inherent limitation to the scope of these findings is that the MTA study allowed for limited medication interventions, with MPH being the most commonly prescribed medication treatment, and a combination of four behavioral strategies as the psychosocial treatment alone intervention. This is in stark contrast to "real-world" clinical settings where providers are likely to use a broader range of FDA approved medications for the treatment of ADHD and families are less likely to be able to meet the time- and cost-intensive demands of the multicomponent psychosocial treatment intervention offered in the MTA. Nevertheless, the findings from the MTA study are important to review as they carry valuable clinical implications.

Effectiveness of Combined Treatment Interventions. When treatment success was stringently defined as a score of 1.0 or lower SNAP-IV summary scores – indicating symptoms of ADHD and ODD as "not at all" or "just a little" present at the conclusion of the intervention – rates of treatment success were higher among individuals in the combination treatment group (68%) compared to those in the medication alone treatment arm (56%) and those in the psychosocial treatment arm (34%) (Swanson et al., 2001). Children in the psychosocial

treatment and combined treatment groups were shown to have comparable outcomes in regards to percentages of positive point system behaviors demonstrated each day (Pelham et al., 2000). However, children in the combined treatment group were shown to perform significantly better than children in the psychosocial treatment group in regards to rule following and good sportsmanship behaviors demonstrated each day, peer ratings of likability, and teacher-reported post-treatment ratings of ADHD severity (Pelham et al., 2000).

Aside from these primary and secondary findings of the MTA study, a recent meta-analysis indicated psychosocial treatment alone (e.g., behavioral therapy) and medication (MPH) alone were both effective in treating symptoms of ADHD in children between the ages of 6 to 12 (Van der Oord et al., 2008). However, the effect size of psychosocial intervention alone (parent-reported $d = 0.87$; teacher-reported $d = 0.75$) was noted to be smaller than that of medication alone (parent-reported $d = 1.53$; teacher-reported $d = 1.83$) or that of the combined treatment intervention (parent-reported $d = 1.89$; teacher-reported $d = 1.77$). These findings suggest psychosocial interventions are less effective than medication or combined treatment interventions in reducing symptomatology of ADHD . Although the effect sizes of the medication alone and the combined treatment interventions were similar, it is important to note that a majority of the combined treatment studies included in this meta-analysis utilized medication alone as the first-line treatment intervention. As such, the treatment gains due to the primary psychopharmacological intervention may have limited the magnitude of supplemental benefit afforded by later psychosocial intervention, which, in combined treatment interventions, was added only once a therapeutic dose of medication had been achieved.

Moderators of Combined Treatment Outcomes. Recent research has indicated certain child and parent characteristics to be moderators of treatment outcomes subsequent to a combined treatment intervention approach. For example, severity of child ADHD symptoms and parental depression at baseline moderate child response rates to combined treatment (Owens et al., 2003). Additionally, in the presence of severe child symptoms of ADHD and parental depression at baseline, below average child IQ is a moderator of poorer treatment response to combined interventions (Owens et al., 2003). There is also evidence to suggest that children with certain psychological comorbidity patterns have differential responses to combined interventions. More specifically, children with two or more psychological comorbidities, as compared to children with one or no psychological comorbidities, demonstrate greater treatment response to combined interventions (Jensen et al., 2001; March et al., 2000). Additionally, children with ADHD+ANX+ODD/CD demonstrate greater treatment response rates to combined treatment interventions compared to children with ADHD+ODD/CD (Jensen et al., 2001).

Sequencing Interventions. A natural and important extension to the evidence base for the use of combined interventions is an investigation of which sequence of medication and psychosocial intervention leads to the most robust treatment outcomes for youth with ADHD. There is evidence to suggest that behavioral intervention alone (e.g., BPT and follow-up behavioral consultation to parents and teachers) can effectively manage symptoms of ADHD, delay the clinical need for adjunctive psychopharmacological intervention, and lead to lower doses of prescribed medications, if they are ultimately clinically indicated (Coles et al., 2019).

There is also evidence to suggest that when behavioral interventions are implemented first, with medication as the adjunctive agent, secondary psychostimulants provide significant added benefit in terms of the child's measure of overall functioning (Pelham et al., 1993). However, when the medication is the first line of treatment, secondary behavioral interventions provide only marginal improvements in primary treatment outcomes (Pelham et al., 1993). It is important to note that these acute outcomes are not indictive of sustained improvements in symptomatology and functioning, and further research is needed to investigate the long-term efficacy of both permutations of combined treatment for youth with ADHD. A more recent investigation of treatment sequencing effects among children between the ages of 5 and 12 concluded that children who began with the behavioral intervention (i.e., BPT plus DRC treatment intervention) demonstrated higher rates of compliance with classroom rules compared to children who began with a stimulant treatment intervention (Pelham et al., 2016). Among those children who did not demonstrate any significant improvement within 8 weeks of their firstly assigned intervention (e.g., medication or behavioral intervention), it was found that an increase in dosing (medication) or intensity (behavioral) of the first-line intervention led to more significant improvements in classroom functioning as compared to maintaining the first-line treatment and augmenting with a secondary intervention (Pelham et al., 2016).

In all, the current literature suggests that psychosocial and pharmacological interventions, two discrete approaches to the treatment of ADHD in childhood, provide complementary clinical benefits to affected youth as they address specific situational impairments that the other intervention cannot directly target (Pelham et al., 1993, 2016). To

this end, one of the most important clinical considerations during treatment planning is the unique features of each patient. Once primary domains of impairment are identified, treatment providers can refer to the evidence-base to guide the selection of treatment interventions best suited to the child's needs.

Trends in Clinical Practice

As the evidence-base for the treatment of ADHD in childhood has evolved over time, so have the established treatment recommendations and the trends in clinical practice. The original treatment recommendations set forth by the AAP indicated stimulant medication and/or behavior therapy as primary treatment interventions for children with ADHD (American Academy of Pediatrics. Subcommittee on Attention-Deficit/Hyperactivity Disorder and Committee on Quality Improvement., 2001). Unsurprisingly, data from clinical care claims for children, ages 2 to 5, with ADHD between the years of 2008 and 2011 suggested approximately 75% of the sample received pharmacological intervention with behavioral intervention only being administered to 55% of children (Visser et al., 2016). In 2011, the AAP's revised treatment recommendations explicitly indicated evidence-based behavioral treatments as a first-line intervention for preschool- and elementary-aged children (AAP, 2011). Although there was a statistically significant increase in the number of families who received behavior therapy services subsequent to the release of newest AAP treatment guidelines, these increased rates of utilization were limited to those families with employer-sponsored insurance to the exclusion of families with Medicaid coverage (Visser et al., 2016). More notably, there continues to be inadequate rates of compliance with evidence-based treatment recommendations that affects

children of various ages, ethnicities, and health insurance statuses. The most recent report of rates of utilized treatment services for youth with ADHD denote medication (66.9%, 95% CI [63.5%, 70.2%]) and school supports (i.e., school-based educational support and/or classroom management) (64.7%, 95% CI [61.4%, 68.0%]) as more commonly utilized interventions compared to psychosocial treatments (i.e., peer interventions, social skills training, CBT, and/or parent training) (32.5%, 95% CI [29.2%, 36.0%]) (Danielson et al., 2018b). Among those receiving psychosocial interventions, data indicated social skills training to be the most commonly implemented psychosocial treatment (39%), with BPT (31%) and cognitive behavioral therapy (CBT; 26%) as the second and third most prevalent psychosocial interventions delivered, respectively. Given that both individual- and organizational-level factors can facilitate or impede the implementation of well-established psychosocial interventions (e.g., BPT, BCM, STP) in clinical care settings, increasing access to evidence-based care (an organizational-level factor) is one pathway to promoting greater rates of compliance with the AAP's treatment recommendations (Danielson et al., 2018b; Wright et al., 2015). To this end, adaptations of established treatments (i.e., altering the structure and/or content of a given treatment protocol to match the needs of a given client population or the limitations of a treatment setting (Eyberg, 2005)) may be critical to increasing access of well-established treatments to families in need.

Statement of the Problem

ADHD in childhood is a striking condition due to its chronic nature (Sibley et al., 2017) and correlation with adverse sequelae across the lifespan such as poor educational and occupational attainment, increased risk for unintentional injuries, elevated health risks, and

early mortality (Barkley et al., 2006; Dalsgaard, 2015; Farone, 2015; Lahey et al., 2016; Nigg, 2012; Owens & Hinshaw, 2016). Approximately two in three children with ADHD have a comorbid mental, emotional, or behavioral disorder (Cuffe et al., 2015). The high rate of psychiatric comorbidity among children with ADHD invariably leads to intensified adversity and functional impairment (Cuffe et al., 2015; Daviss, 2008; Loe & Feldman, 2007; Owens & Hinshaw, 2016; Riglin et al., 2020). Two transitional periods in youth development have been identified as "sensitive periods" for therapeutic intervention and effective management of ADHD: the transition from childhood to adolescence, and the transition from adolescence to young adulthood (Turgay et al., 2012). It is of significant clinical importance to assess for and treat ADHD in childhood in order to mitigate the scope and duration of negative sequelae associated with this chronic disorder.

The current evidence-base for the treatment of ADHD in childhood indicates three categories of treatment approaches: psychosocial intervention (e.g., BPT, BCM, STP), psychopharmacological intervention (i.e., stimulant medication), and combined intervention (i.e., concurrent use of psychosocial and pharmacological treatments) (Brown et al., 2008). Of these three treatment approaches, psychosocial intervention is recommended as the first line of treatment for preschool and elementary age children with ADHD (Evans et al., 2018; Wolraich et al., 2019). The STP, one of the three well-established psychosocial interventions for ADHD in childhood, is an intensive, multimodal behavioral treatment which utilizes a reward and response-cost system, a time out protocol, and a DRC intervention to shape and reinforce prosocial and on-task behaviors and punish maladaptive behaviors common to children with

ADHD and ODD (Fabiano et al., 2014; Pelham & Hoza, 1996). Over the span of 8 weeks, the STP provides children with 280 hours of direct intervention and parents with 8 hours of group-based BPT (Fabiano et al., 2014; Pelham & Hoza, 1996). The STP facilitates improvements in social, recreational, and academic functioning for children with ADHD regardless of their gender, age, and psychiatric comorbidities (Coles et al., 2005). Despite the demonstrated effectiveness of the STP, this treatment modality is inaccessible to a majority of families in need: the STP is currently offered at only fifteen sites across the US, and the 8-week program often involves substantial costs to families who are interested in participating.

In an effort to increase the accessibility of the well-established STP intervention, recent research has sought to examine adaptations to the standard STP protocol that are more feasible for clinical care settings to implement and more acceptable for families with limited time- and monetary-based resources. There is evidence to suggest that 2- and 3-week long STP programs, that maintain the daily schedule established in the standard STP, yield significant parent-reported reductions in symptoms of ADHD and ODD (Yamashita et al., 2010; Yamashita et al., 2011). However, these brief STP interventions did not lead to significant improvements in peer relationships and prosocial behaviors (Yamashita et al., 2010). In contrast, a 6-week adaptation of the STP protocol has demonstrated feasibility, acceptability, and efficacy in reducing the severity of children's internalizing, externalizing, and attention problems (Tannenbaum et al., 2019). Despite preliminary evidence and support for implementation of this adapted STP, further investigation is needed to understand the overall effectiveness of this treatment model in terms of clinically meaningful and interpretable outcomes and to identify

the subpopulations of children with particularly strong or poor response to this specific intervention.

To address this limitation in the literature, this study proposes to: (1) investigate the effectiveness of the 6-week adaptation of the STP model by examining behavioral improvement, as defined by (a) change in frequency of behaviors across standard STP point system categories, (b) change in percentage of daily report card target behaviors achieved, and (c) differences in pre- and post-treatment parent-reported measures of child inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations; and (2) to investigate whether children who receive concurrent pharmacotherapy experience greater benefits from the 6-week STP. A significant strength of this investigation will be the inclusion of children with one or more psychiatric comorbidities and/or a history of psychopharmacological treatment, as these are common clinical characteristics of children who seek treatment in community care settings.

Research Questions and Hypotheses

Research Question 1

Do children who complete the 6-week STP show improvements in parent-reported outcomes, including: inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations?

Hypothesis 1. Parent-reported scores of inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations will decrease following the completion of the 6-week STP.

Research Question 2

Do children who complete the 6-week STP show improvements in frequency of positive behaviors, including: answering attention questions, following activity rules, complying with counselor commands, contributing to group discussions, ignoring a negative stimulus, helping a peer, and sharing with a peer?

Hypothesis 2. The frequency of behaviors within each positive behavior category will increase throughout the 6-week STP.

Research Question 3

Do children who complete the 6-week STP show decreases in frequency of negative behaviors, including: interrupting, whining, swearing, teasing, verbal abuse, leaving the activity area, intentional aggression, and intentional destruction?

Hypothesis 3. The frequency of behaviors within each negative behavior category will decrease throughout the 6-week STP.

Research Question 4

Do children who complete the 6-week STP show improvements in academic learning center outcomes, as demonstrated by total points earned for: following learning center rules (Behavior score), completing homework assignments and for completing assignments with 80% or higher accuracy (Academic score), and other positive behaviors (bonus points)?

Hypothesis 4. The points earned for positive behaviors (Behavior score plus bonus points) and homework completion and accuracy (Academic score) will increase throughout the 6-week STP.

Research Question 5

Do children who complete the 6-week STP show improvements in the percentage of DRC target behaviors achieved?

Hypothesis 5. The percentage of DRC target behaviors achieved will increase throughout the 6-week STP.

Chapter III: Methods

Participants

The sample includes 52 children, ages 6 to 12 years, and their caregivers who completed the 2019 Camp Baker program. Among these 52 children, 88.5% (n = 46) were male and 65.4% (n = 34) were Caucasian. The majority of youth (61.5%, n = 32) were diagnosed with at least one comorbid mental health disorder in addition to ADHD. Common comorbid disorders included anxiety (20.4%, n = 10), depression (10.2%, n = 5), ODD (8.2%, n = 4), and Autism Spectrum Disorder (14.3%, n = 7). A small subset of the participant sample (9.6%, n = 5) did not have a diagnosis of ADHD. At the time of admission, 82.7% (n = 43) of youth were receiving pharmacotherapy for one or more of their psychological disorders. Within this subset of participants, 44.2% (n = 23) were prescribed an FDA-approved medication for ADHD (ADHD Only), 17.3% (n = 9) were prescribed an FDA-approved medication for ADHD in addition to one or more psychotropics, with indication(s) other than ADHD (ADHD + Other), and 21.2% (n = 11) were prescribed one or more psychotropics, with indication(s) other than ADHD (Other). A majority of families (73.1%, n = 38) reported a household income of at least $100,000. Table 1 below presents descriptive characteristics for baseline demographic and clinical characteristics of the study sample.

Table 1

Descriptive Statistics: Baseline Characteristics

	Total Sample (N = 52)	ADHD only (n = 23)	ADHD + Other (n = 9)	Other (n = 11)	None (n = 9)	p
Age (Mean, *SD*)	8.29, *1.74*	8.17, *1.44*	9.11, *1.36*	8.64, *2.34*	7.33, *1.73*	.133

Table 1 (continued)

Gender						.245
Male, *n* (%)	46 (88.46%)	19 (82.61%)	7 (77.78%)	11 (100%)	9 (100%)	
Female, *n* (%)	6 (11.54%)	4 (17.39%)	2 (22.22%)	0 (0%)	0 (0%)	
Race, *n* (%)						.644
Caucasian	34 (79.07%)	14 (77.78%)	6 (75%)	9 (90.00%)	5 (71.43%)	
African American	3 (6.98%)	2 (11.11%)	0 (0%)	1 (10.00%)	0 (0%)	
Asian	2 (4.65%)	1 (5.56%)	0 (0%)	0 (0%)	1 (14.29%)	
Biracial	4 (9.30%)	1 (5.56%)	2 (25%)	0 (0%)	1 (14.29%)	
(Missing)	9	5	1	1	2	
Ethnicity, *n* (%)						.659
Hispanic	10 (50%)	5 (71%)	1 (33%)	2 (40%)	2 (40%)	
Non-Hispanic	10 (50%)	2 (29%)	2 (67%)	3 (60%)	3 (60%)	
(Missing)	32	16	6	6	4	
Diagnosis Pattern, *n* (%)						.117
ADHD only	15 (28.85%)	10 (43.48%)	2 (22.22%)	1 (9.09%)	2 (22.22%)	
ADHD + Anxiety	14 (26.92%)	4 (17.39%)	4 (44.44%)	4 (36.36%)	2 (22.22%)	
ADHD + Other	18 (34.62%)	9 (39.13%)	3 (33.33%)	3 (27.27%)	3 (33.33%)	
Other	5 (9.62%)	0 (0%)	0 (0%)	3 (27.27%)	2 (22.22%)	
Grade, *n* (%)						.213
K	7 (13.73%)	1 (4.35%)	0 (0%)	2 (18.18%)	4 (44.44%)	
1st	8 (15.69%)	5 (21.74%)	0 (0%)	2 (18.18%)	1 (11.11%)	
2nd	12 (23.53%)	7 (30.43%)	1 (12.50%)	2 (18.18%)	2 (22.22%)	
3rd	10 (19.61%)	5 (21.74%)	3 (37.50%)	1 (9.09%)	1 (11.11%)	
4th	9 (17.65%)	4 (17.39%)	3 (37.50%)	1 (9.09%)	1 (11.11%)	
5th	3 (5.88%)	0 (0%)	1 (12.50%)	2 (18.18%)	0 (0%)	
6th	2 (3.92%)	1 (4.35%)	0 (0%)	1 (9.09%)	0 (0%)	
Total Family Income (Mean, *SD*)	$239,647, *216,034*	$265,744, *272,622*	$145,325, *118,331*	$253,418, *124,497*	$250,000, *217,721*	.561

Table 1 (continued)

Total Family Income, *n* (%)					
< $50,000	8 (15.38%)	5 (21.74%)	2 (22.22%)	1 (9.09%)	0 (0%)
$50,000 - $74,999	1 (1.92%)	0 (0%)	0 (0%)	0 (0%)	1 (11.11%)
$75,000 - $99,999	4 (7.69%)	2 (8.70%)	1 (11.11%)	1 (9.09%)	4 (7.69%)
$100,000 - $199,000	17 (32.69%)	7 (30.43%)	4 (44.44%)	1 (9.09%)	5 (55.56%)
$200,000 - $299,999	7 (13.46%)	2 (8.70%)	1 (11.11%)	3 (27.27%)	1 (11.11%)
$300,000 - $499,999	10 (19.23%)	3 (13.04%)	1 (11.11%)	5 (45.45%)	1 (11.11%)
> $499,999	5 (9.62%)	4 (17.39%)	0 (0%)	0 (0%)	1 (11.11%)

Procedures

Recruitment

Families were referred to Camp Baker by Judge Baker Children's Center (JBCC) staff and local mental health professionals in the Boston area. The primary referral sources include pediatricians, school psychologists and guidance counselors, and word of mouth from prior participants. Some families inquired about Camp Baker after reviewing brochures on the program website. Camp Baker also pays for advertising on social media, in local parenting papers, and at local camp fairs.

Eligibility Screening

In the first step of the Camp Baker application process, interested families were asked to submit a non-refundable application fee and complete a questionnaire that solicited the following: child information (age, gender, developmental history, medical history, previous psychological evaluations, diagnoses, and treatments), caregiver demographics, and family information (parents' marital status, total household income, family medical history, number of

individuals living in the home). In the second step of the application process, interested families were asked to return parent- and teacher-reported rating scales and to submit copies of their child's most recent individualized education program (IEP) or 504 plan and neuropsychological or psychological evaluation, if applicable. The parent-reported rating scales in this second step of the application process included the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001), the Conners 3rd Edition – Parent Short Form (Conners 3-P(S); Conners, 2008), and the Parenting Stress Index (PSI; Abidin, 1995). Caregivers were responsible for instructing the child's current teacher to complete the following teacher-reported rating scales: the Teacher Report Form (TRF; Achenbach & Rescorla, 2001) and the Conners 3rd Edition – Teacher Short Form (Conners 3-T(S); Conners, 2008).

Completed applications were then reviewed by the Associate Director of Camp Baker (PsyD) and a Program Assistant (M.Ed.). As the final step of the application process, the Associate Director of Camp Baker conducted phone interviews with caregivers to further assess the child's psychological history and eligibility. Children were admitted to the program if they had a previous diagnosis of ADHD and/or demonstrated significant attention, externalizing, or internalizing symptoms, as determined by the Brief Problem Monitor (BPM). Children were excluded if they carried a diagnosis of posttraumatic stress disorder (PTSD) or bipolar disorder, or if caregivers and/or teachers reported a history of daily physical management, severe physical aggression, or severe property damage. Children were allowed to participate in the summer program if they were previously or currently on medication. These inclusion and

exclusion criteria were in place to maximize accessibility to children and families in need of intensive behavioral treatment and to reflect common practice in community care settings.

Staff Training and Supervision

In addition to providing intensive evidence-based intervention to children and families, Camp Baker seeks to provide undergraduate and graduate student interns with rigorous training and clinical experience. Each Camp Baker staff member (i.e., Graduate Counselors, Undergraduate Counselors, Special Education Teachers, Teachers-in-Training, Teacher Aides) received a copy of the Camp Baker treatment manual. Camp Baker leadership staff led mandatory training of summer staff over a period of 6 days prior to the start of Camp Baker. Staff were required to read the Camp Baker treatment manual prior to attending the training and completed pre-training quizzes to assess their baseline knowledge of the STP. During the training period, staff received intensive training on the STP point system and behavior management protocol, both of which were detailed in the Camp Baker treatment manual. Comprehension quizzes were completed throughout staff training as a learning tool so that staff were aware of what knowledge they had acquired and areas for continued growth. A strong emphasis was placed on learning how to implement the point system with a high degree of reliability. At the conclusion of the summer program, Camp Baker staff members received a small stipend.

Throughout the 6-week STP, Camp Baker leadership ensured treatment integrity by providing live-supervision to staff members, thereby identifying any treatment protocol violations and providing immediate corrective feedback. In addition, all staff members regularly

participated in scheduled supervision meetings. These meetings allowed counselors the opportunity to elicit guidance as to which therapeutic strategies or approaches would be recommended in response to certain child behaviors or group dynamics, while maintaining fidelity to the treatment protocol. Supervision meetings also provided an opportunity for the Director to provide ongoing feedback to staff members about their treatment implementation skills. The Associate Director completed fidelity checks on a weekly basis. Inter-rater reliability for the STP point system was not formally assessed.

Camp Baker

Camp Baker is one of the fifteen Summer Treatment Programs (STP) offered across the United States. Unlike these other programs, however, Camp Baker is an adaptation of the original STP treatment protocol (Pelham & Hoza, 1996; Fabiano et al., 2014). Briefly, Camp Baker's adaptations to the traditional STP protocol include: (1) reducing the length of the program from 8 weeks to 6 weeks, (2) implementing the behavioral parent training curriculum from the Modular Approach to Treatment for Children with Anxiety, Depression, Trauma or Conduct Problems protocol (MATCH-ADTC; Choprita & Weisz, 2009), (3) delivering the treatment in a traditional camp setting, providing a more authentic camp experience in an effort to reduce stigma, (4) providing an expanded selection of games played in the recreational setting, and (5) providing treatment reports and consultation services to support the consistent use of strategies and service delivery across settings. These adaptations to the structure and content of the previously established STP model were intended to increase the accessibility of this well-established treatment to children and families in need of services.

Although Camp Baker offered 70 less intervention hours than the traditional STP (Pelham & Hoza, 1996), children and families received a total of 210 hours of direct intervention, equivalent to approximately 4 years' worth of standard outpatient care which would otherwise be delivered in the form of one-hour, once-weekly sessions. Thus, Camp Baker was able to provide an intensive treatment while requiring two less weeks of cost for families, in terms of both time and financial resources. Additionally, Camp Baker implemented a more targeted parent training curriculum: MATCH-ADTC. In these weekly, group parent training sessions, parents received didactics on child development and disruptive behavior, strategies to enhance parent-child interactions, and guidance on implementing effective contingency management systems at home to improve child behavior. In contrast to the Community Parent Education Program protocol (COPE; Cunningham, 2005) implemented in the traditional STP, parents were not asked to watch videotaped vignettes of problematic situations or engage in small group problem solving discussions. Rather, Camp Baker parent training sessions each involved one didactic on a skill or topic of the week, one role-play, where a licensed psychologist modeled the use of the parenting strategy introduced that week, and a time for question and answers so that parents could receive further instructions and guidance on how to effectively implement the recommended parenting strategy.

Therapeutic Milieu. The STP intervention implemented at Camp Baker involved seven hours of programming, five days a week, across six weeks during the summer. Daily activities at Camp Baker included morning meeting, swim lessons, academics, art, skill drills, sports games, and reward recess. Table 2, below, presents the daily schedule and further details about

each activity. Unless noted otherwise, points were called and recorded during each activity period.

Table 2

Daily Schedule of Activities

Time	Number of Intervals	Activity	Description
8:00 AM	*Not Applicable*	Bus to Camp	Parents dropped children off at one of two bus stops: one in Boston, MA and one in Westwood, MA. Counselors then rode with children to Camp Hale Reservation in Westwood, MA (site of Camp Baker). Points were called but not recorded during this activity period.
8:45 AM	1	Morning Meeting	For the first 15 minutes of the day, all groups engaged in a group discussion in which counselors reviewed children's progress towards weekly goals, presented daily awards to four children, and presented a brief social skills training module. Daily award recipients were determined by previous day points tallies in areas of: most points, most improved, social skills, and sportsmanship. Counselors presented social skills training adapted from Pelham & Bender (1982) and Pelham et al., (1988). Social skills modules included: communication, cooperation, validation, and participation.
9:00 AM	1	Transition	
9:10 AM	3	Swim Lessons	After completing a swim test on the first day of camp, children participated in structured swimming lessons matching their ability levels.
9:55 AM	1	Transition	
10:05 AM	*Not Applicable*	Snack	Points were called but not recorded during this activity period.
10:15 AM	4	Academics	Academic Learning Center staff members (Special Education Teachers, Teachers-in-Training, Teacher Aides) led children in the completion of academic activities and assignments in a classroom setting.
10:55 AM	1	Transition	

Table 2 (continued)

11:05 AM	4	Skill Drill	Counselors taught and coached children in skills pertinent to the sports game of the day (i.e., soccer, kickball, baseball). During the first interval, counselors led a pre-activity discussion to discuss the rules and procedures of the upcoming drills. During the final activity interval, counselors led a post-activity discussion to discuss any problems that occurred, progress that was made, and examples of how the social skill of the day was used during the skills drill activity.
11:45 AM	1	Transition	
11:55 AM	*Not Applicable*	Lunch	Points were called but not recorded during this activity period. Counselors scored children's AM DRC intervals and reviewed scores with children, providing feedback on areas of improvement for PM intervals so as to meet the child's DRC goal(s) for the day.
12:10 PM	*Not Applicable*	Reward Recess	Children who scored 50% or higher on their AM DRC intervals were awarded free recess. Children who did not meet at least 50% of their AM DRC goals were escorted to a designated sit-out area. For all children, points were called but not recorded during this activity period.
12:20 PM	1	Transition	
12:30 PM	4	Sports Game	Children engaged in game play, practicing their sports skills in a naturalistic setting. During the first interval, counselors led a pre-activity discussion to discuss the rules and procedures of the game. During the final activity interval, counselors led a post-activity discussion to discuss any problems that occurred, progress that was made, and examples of how the social skill of the day was used during the sports game.
1:10 PM	1	Transition	
1:20 PM	4	Art	Children engaged in art activities in a classroom setting.
2:00 PM	1	Transition	

Table 2 (continued)

2:10 PM	*Not Applicable*	Free Swim	Points were called but not recorded during this activity period.
2:55 PM	*Not Applicable*	Transition	Points were called but not recorded during this activity period.
3:05 PM	*Not Applicable*	Reward Recess	Points were called but not recorded during this activity period.
3:30 PM	*Not Applicable*	Bus Home	Points were called but not recorded during this activity period.

Children were placed in one of four age-matched groups: young (ages 6 – 7), mid-young (ages 7 – 8), mid-old (ages 8 – 10), and old (ages 10 – 12). However, there were two exceptions: at the Associate Director's discretion, one 7-year-old boy and one 11-year-old girl were placed in the mid-old group. Each group of children was staffed with two graduate counselors and three to four undergraduate counselors.

Contingency Management. In keeping with the traditional STP model, Camp Baker implemented a behavioral contingency management protocol whereby principles of reinforcement and punishment were applied to child behavior throughout the day (Pelham & Hoza, 1996; Fabiano et al., 2014). Throughout the day, social reinforcement, including labeled praises and public acknowledgement in the form of daily awards, was also utilized to reinforce positive child behaviors.

Standard STP Point System. Camp Baker counselors implemented the standard STP point system during every activity period during the camp day with the exception of academic learning center (ALC; 45 minutes), lunch time (25 minutes), reward recess (10 minutes), and free swim (35 minutes). The STP point system includes ten categories of appropriate ("positive") behaviors, for which children can earn points, and fifteen categories of inappropriate

("negative") behaviors, for which children can lose points (Appendix A). For each behavior a child demonstrated during the camp day, a Counselor would decide if the behavior met the operational definition of one of the twenty five behavior categories. If the behavior met criteria for one or more point system categories, the Counselor would classify the behavior according to the Rules for Classifying Point System Behaviors (Synn et al., 2019). Each week, children were able to exchange the points they had earned for tangible rewards / prizes.

ALC Point System. During the children's time in ALC (45 minute interval), a separate reward and response cost system was implemented by the teaching staff. As children completed seatwork assignments, two ALC staff members circulated through the room to correct assignments at children's desks, provide immediate feedback to children regarding their accuracy, and to call out points earned for academic performance and points lost for rule violations. Meanwhile, one ALC staff member graded homework assignments and assisted with recording points earned and lost on the public point board.

Daily Report Card (DRC). For some children, a point system alone is not sufficient to produce the desired changes in one or more category of behavior. The DRC is recognized as a best practice for youth with ADHD (Fabiano et al., 2010; Volpe & Fabiano, 2013). In conjunction with the STP and ALC point systems, the implementation of DRCs allowed for a more individualized and targeted approach to treatment. Incorporating information from the intake interview, the child's history of treatment, observations of the child, and the child's demonstrated behaviors during the first week at Camp Baker, the child's counselor and clinical supervisors collaboratively developed target behaviors for each child's DRC. Each child was

assigned 3-5 DRC target behaviors which included one or more goals related to group activities, the classroom, or social behaviors (Appendix B). Possible target behaviors included ones such as: "Following Activity Rules for 75% of all intervals" (in this case, the target behavior seeks to augment the power of the STP point system) and "Take a break appropriately when prompted by staff" (in this case, the target behavior is not one that is not included in the STP point system, but rather a socially valid target behavior which serves to increase the child's social functioning in a specific manner the STP point system does not specify). Children's daily DRC percentages were reviewed weekly during supervision meetings. Leadership staff and Group Counselors collaborated on revisions to a child's DRC target behaviors, as needed, to ensure goals were achievable and of a gradually increasing difficulty over time.

At the end of each camp day, parents received a copy of their child's scored DRC and were counseled on whether to reward their child for reaching at least 75% of their DRC target behaviors or whether to withdraw a privilege at home as a consequence for a child meeting less than 50% of their DRC target behaviors. Thus, the DRC provided an additional layer of contingency management.

Parent Involvement. Once weekly, parents attended a behavioral parent training group session led by the Associate Director of Camp Baker. Over the six week period, parents received education on child development and disruptive behavior in addition to skills training on how to establish and maintain positive parent-child interactions, generate effective instructions for their child, and consistently implement a home-based reward and response cost system. Throughout the subsequent school year, parents attended monthly group sessions in which

parents received additional didactics and skills training to help maintain and extend their children's summer treatment gains.

Data Collection. Paper point sheets were used to immediately record points awarded for positive behaviors and points lost for negative behaviors throughout the day. Paper DRCs were scored twice daily, once in the AM interval (during Lunch) and once in the PM interval (Free Recess, prior to children returning home). Staff were responsible for entering data from paper point sheets and DRCs into an electronic data set prior to the end of each camp day. During ALC, staff maintained a public point board on which tallies of children's rule violations were recorded and points earned for completing seatwork assignments was recorded. ALC staff were responsible for entering data from the public point board into the aforementioned electronic data set at the conclusion of each group's ALC period. Leadership staff conducted periodic audits of all data entered, cross-referencing filed paper records with electronic dataset entries to verify accuracy of data entry.

Measures

Behavioral Data

The present study seeks to investigate specific categories of positive and negative behaviors which match the most relevant target behaviors for children with ADHD, ODD, and related behavioral disorders of childhood. These specific categories and their operational definitions are detailed below (Pelham et al., 2017; Synn et al., 2019). Unless otherwise noted, weekly frequency counts of a given behavior were used for the statistical analyses.

Positive Point System Behaviors.

Following Activity Rules. Every activity period at Camp Baker was introduced to children by a review of the rules specific to that activity. Throughout the activity period, counselors noted any violations of activity rules. Children earned points for each interval of the activity period in which they did not engage in any behaviors which violated the activity rules. If a child did not obtain points for following activity rules in one of the intervals, the subsequent interval of the activity allowed the child a new opportunity to earn points.

Attention. During three activities each day, Counselors assessed children's attention to the ongoing activity by asking a specific question related to instructions or information a counselor recently verbalized. Children were asked attention questions of varying difficulty, with questions ranging from one to three parts (i.e., What is a rule of this activity? What is the last idea that was contributed and who contributed it? What were the last two rules that were contributed and who contributed them?). The difficulty level of the attention question assigned to each child was determined during Counselors' group supervision with Leadership Staff. Children earned points for answering all parts of their assigned attention question correctly. Irrespective of the level of difficulty assigned to a child, the standard point value (Appendix A) for answering an attention question was awarded each time a child demonstrated accurate recall. In this study, the average percentage of questions answered correctly across each week was used for the statistical analysis.

Complying with a command. Children earned points for compliance when they initiated or ceased to engage in a behavior within ten seconds, or within a specified time, of a direct

command from a Counselor. Counselors evaluated compliance for both commands directed to an individual child and for commands directed toward a group of children.

Contributing to discussion. During social skills training modules of the morning meeting, pre- and post-activity discussions, ALC, and any other miscellaneous group discussions, children earned points for verbally providing a task-related, non-redundant statement. In social skills training sessions, qualifying contributions include: presenting definitions or role-playing examples of appropriate and inappropriate social behaviors, stating importance of learning and using social skills, and describing a possible outcome of using a social skill. During activity discussions, qualifying contributions include: defining the rules of the activity, describing how points can be earned or lost during the activity, suggesting how social skills can be demonstrated during the activity, discussing problems that occurred or progress that was made during the activity, and presenting ideas for improving the group's success in the activity in the future. Qualifying statements in miscellaneous group discussions included: responding constructively to a Counselor's inquiry, asking a constructive question, and presenting ideas relevant to the discussion.

Ignoring a negative stimulus. Children earned points when they did not engage in an observable negative response (i.e., interruption, whining, swearing, teasing) to any verbal or nonverbal behavior from another child (i.e., teasing, intentional aggression) that would otherwise elicit annoyance or distress in the recipient child. If a child continuously directed negative behaviors towards other children, those recipient children who ignored the behavior were awarded points for ignoring once every minute.

Helping a peer. Children were awarded points for helping when they voluntarily provided a peer assistance or aid that: was relevant to the ongoing activity, offered in a manner that did not disrupt the ongoing activity, did not meet criteria for any negative behavior category, and was accepted by the peer. Of note, Counselors did not award points for helping when a child demonstrated a behavior that was a necessary component of a game (i.e., passing to a team member).

Sharing with a peer. Children were awarded points for sharing when they voluntarily provided a peer with a personal possession, privilege, or material and so long as the sharing behavior: was relevant to the ongoing activity, offered in a manner that did not disrupt the ongoing activity, did not meet criteria for any negative behavior category, and was accepted by the peer. Counselors did not award points for sharing when a child demonstrated a behavior that was a necessary component of a game.

Negative Point System Behaviors.

Interrupting. Children lost points each time they engaged in a verbal or nonverbal behavior which intruded on the activity (i.e., in a group setting, a child engages in a behavior which results in two or more peers engaging in behaviors, for at least two seconds, which are incompatible with the ongoing activity) or others' conversation (i.e., in a dyadic interaction, a child engages in a disruptive behavior at the same time another person is talking).

Whining. Children lost points each time they demonstrated a verbal or nonverbal behavior that inappropriately expressed discomfort, dissatisfaction, or resentment through content, gesture, or tone of voice. Children who demonstrated appropriate expressions of

discomfort or dissatisfaction (i.e., expressed in a neutral tone of voice and normal pitch and intensity, not antagonistic in content, and not accompanied by unnecessary or negative gestures) did not lost points for whining.

Swearing. Children lost points each time they demonstrated a verbal or nonverbal behavior, regardless of tone or intensity, that would typically be regarded as profane, obscene, or offensive and that was not directed toward an individual who could see or hear it. If a child engaged in swearing, the child's Lead Counselor informed parents of the behavior during check-out.

Teasing. If a child directed a negative communication toward a peer who was identifiable as the intended recipient and who could see or hear the negative communication, the child lost points for teasing. Negative communications were defined as either a derogatory name or any other verbal or nonverbal behavior which would typically elicit a clear behavioral indication of annoyance or distress from the intended recipient.

Verbal abuse. Children lost points for directing a negative communication towards a Camp Baker staff member (i.e., leadership staff, Counselor, Teacher) who was identifiable as the intended recipient and who could see or hear the negative communication.

Leaving the activity area. At the beginning of each activity period, Counselors defined the boundaries of the activity area. Children lost points for leaving the designated activity area without permission from a Camp Baker staff member. Children continued to lose points for each minute they remained outside of the activity area.

Intentional aggression. Children lost points when they intentionally performed a physical behavior that: would typically produce physical injury or pain to another, or intruded on another by inappropriately restricting their freedom of movement. Children also lost points for intentional aggression if they verbalized threats or used racially charged language. Of note, the staff member who observed the child's act of aggression was responsible for determining intent of the behavior. Children who engaged in an act of intentional aggression were immediately assigned a time out.

Intentional destruction. Children lost points when they intentionally performed a physical behavior that: destroyed or damaged an object, defaced an object's surface, or otherwise altered an object so that its value or usefulness was substantially impaired or at least reduced temporarily. Of note, the staff member who observed the child's act of destruction was responsible for determining intent of the behavior. Children who engaged in an act of intentional destruction were immediately assigned a time out. Behaviors that appeared to be the result of clumsiness, lack of skill, or inattention were considered unintentional.

ALC Point System Behaviors. The present study seeks to investigate performance in the academic setting, as this is a commonly reported presenting concern, source of distress, and area of impairment. The ALC total score is a summation of two categories of performance (behavior, academic). These categories, and the way in which point totals were calculated, are detailed below. In this study, the weekly ALC total scores were used for the statistical analyses.

Behavior. At the beginning of each ALC period, staff reviewed the learning center rules posted on the board in the front of the classroom. Children were also awarded 100 points for

behavior at the outset of each ALC period. Children lost 10 points for each rule violation. The maximum total behavior score was 100, and the total behavior score could not be less than zero. As such, staff members recorded a total behavior score of zero for children who violated more than 10 rules in one ALC period. The rules specific to ALC, and their operational definitions, are detailed below.

Be respectful of others. The child behaves in a manner that typically does not offend, disrupt, or harm self or other persons. This includes behaviors that would be categorized as cheating or behaviors that meet the criteria for intentional aggression, intentional destruction, stealing, lying, verbal abuse, teasing, or swearing.

Obey adults. When an ALC staff member provides a direct command, the child exhibits or ceases to exhibit a behavior, as specified within the original command, within ten seconds of the direct command or within the time specified by the ALC staff member. Compliance to commands were assessed for both commands directed to an individual child and for commands directed toward a group of children. If the child makes a reasonable and appropriate effort to comply with the command but does not complete the desired behavior within ten seconds, the child did not lost points for noncompliance.

Work quietly. The child does not exhibit any verbal or nonverbal behavior that intrudes into the activity or conversation of others. In a group setting of seatwork (three or more people), behaviors were defined as intrusive if they would typically result in two or more people other than the child in engaging in behaviors, for at least two seconds, that are incompatible with the ongoing activity. In dyadic interactions during seatwork, behaviors were defined as

intrusive if they occurred at the same time another person was speaking. Children were marked for violating this rule if they engaged in a behavior that disrupted the class, causing the teacher to stop or pause the class activity.

Use materials and possessions appropriately. The child uses materials and possessions only for the purpose and in the manner for which they were designed and intended.

Remain in your assigned seat or area. The child remains in his or her seat or area until he or she is given permission to move to another location. “In seat” is further defined as sitting in the chair, facing forward, with all four chair legs on the floor. “Assigned area” refers to the area that is designated by the teacher at a particular point in time.

Raise hand to speak or to ask for help. The child raises his or her hand and waits for acknowledgement from an ALC staff member before speaking, unless otherwise specified.

Stay on task. The child attends to the current assigned task. Attention is indicated by the child: looking at or manipulating objects or materials on the child’s desk that are necessary for completing the task; looking at the blackboard, computer screen, or other location where materials related to the task are displayed; looking at an instructor who is in the process of providing instruction regarding the task; looking at any object or place to which the child has been directed by the instructor; looking at a peer who has been asked a question by the instructor during a group lesson; performing motor activity as required by the task; or performing a motor activity for the purpose of preparing for or finishing a current assigned activity. Children were marked for violating this rule if they did not attend to the assigned task

for at least 10 seconds. If a child was continuously off task, an ALC staff member informed the child of a rule violate once per minute.

Bonus Points. Children were awarded additional points for positive behaviors from the standard STP point system (following activity rules, complying with counselor command, contributing to group discussion, ignoring a negative stimulus, helping a peer, and sharing with a peer) demonstrated during ALC. This allowed for greater opportunities to provide reinforcement and help further shape desired positive behaviors. In this study, bonus points awarded in ALC were included in the total Behavior score.

Academic. Each day, children were assigned seatwork assignments (30 minutes of the ALC period). These assignments emulated traditional schoolwork by providing children independent practice in completing paper-and-pencil assignments in the areas of Math, Reading, and Language. Children's teachers were asked to provide individual assignments for the child to complete across each of these three content areas, as the teachers were uniquely able to assess the child's ability levels and specific areas for growth. From these teacher-curated sets of assignments, ALC staff then assigned children one assignment for each of the learning areas.

Children earned 25 points for completing each of their three assignments and an additional 25 points for each assignment completed with 80% or higher accuracy. Children also earned 25 points for returning completed homework and an additional 25 points if the assignment was completed with 80% or higher accuracy. Daily total academic scores were obtained by summing the points earned for completion and for accuracy, on both seatwork and

homework assignments. Homework was not assigned over the weekend. Thus, children were able to earn a maximum of 150 total academic points on Monday and a maximum of 200 total academic points on Tuesday, Wednesday, and Thursday.

DRC Target Behaviors. DRCs were instituted beginning Week 2; as such, there are 5 weeks' worth of DRC data for each child. Each child was assigned a Graduate or Undergraduate Counselor who was responsible for completing the DRC. These assignments were generated by the Associate Director after reviewing counselor skill level and the clinical complexity of the child's presentation. To complete a child's DRC, counselors were required to: review daily STP point system data, confer with ALC staff or reviewing the ALC point system data, and/or refer to child self-reports for any interpersonal target behaviors (i.e., for a target behavior of "Ask a new friend a question", the Counselor would confirm with the friend / peer that the question was asked), as needed. Once in the morning (first 5 intervals of the day; "AM") and once again in the afternoon (final 5 intervals of the day; "PM"), the Counselor marked a target behavior as having been met (YES), not having been met (NO), or the child not having had the opportunity to complete the target behavior (N/A). Daily DRC percentages were then calculated by summing the total number of YES marks and dividing it by the total number of intervals with opportunity (YES + NO). The present study seeks to investigate whether children who complete the adapted STP achieved higher DRC percentages over time. As such, average weekly DRC percentages were examined across five timepoints.

Parent Ratings

Conners 3rd Edition – Parent Short Form *(Conners 3-P(S);* Conners, 2008).

The Conners 3-P(S) consists of 43-items across six content areas: Inattention, Hyperactivity/Impulsivity, Learning Problems, Executive Functioning, Defiance/Aggression, and Peer Relations. Parents are asked to consider child functioning over the past month and to then rate items using a four-point Likert scale (0 = not true at all; 3 = very much true). The Conners 3-P(S) was adapted from the 110-item Conners 3rd Edition – Parent Report. The Conners 3-P(S) was designed to provide a more brief assessment of the presence of DSM-5 ADHD symptoms and common comorbid problems among youth ages 6 to 18 years old. This measure is also commonly used to monitor severity of symptoms across a period of clinical intervention. Internal consistency reliability, as measured by Cronbach's alpha, was moderate to high for each of the six subscales: Inattention (α = 0.91), Hyperactivity/Impulsivity (0.90), Learning Problems (0.80), Executive Functioning (0.82), Defiance/Aggression (0.89) and Peer Relations (0.88) (Gomez & Vance, 2018). The Conners 3-P(S) has also demonstrated 2- to 4-week test-retest reliability, as measured by Pearson's *r*, ranging from .71 to .98 ($p < .001$; Rzepa & Marocco, 2008). In this study, Conners 3-P(S) scores across each of the six subscales, obtained at pre- and post-treatment, were used for the proposed analyses.

Brief Problem Monitor *(BPM; Achenbach et al., 2011).* The BPM is adapted from 113-item Child Behavior Checklist (CBCL), one of the most widely utilized measures of child functioning. The BPM consists of 19-items across three subscales: Internalizing Problems, Externalizing Problems, and Attention Problems. The BPM has demonstrated convergent validity with the CBCL, as measured by Pearson correlations, with respect to total scores (r = .95) as well as scores for the Internalizing (.86), Externalizing (.93) and the Attention (.97) subscales (Piper et

al., 2014). The BPM has also demonstrated high internal consistency reliability ($\alpha = 0.91$) and satisfactory internal consistency in regards to the Internalizing (0.79), Externalizing (0.86) and Attention (0.87) subscales. In this study, BPM total scores, obtained at pre-treatment, were used as a covariate in the proposed analyses.

Reliable Change Index

The Reliable Change Index (RCI; Jacobson & Truax, 1991) is a psychometric criterion used to evaluate statistical and clinical significance of a change over time. It is unique in that it takes into account the scores at two time points (e.g., pre- and post-treatment) and the standard error of difference between them. This standard error of difference, in turn, takes into account the standard deviation of the pre-treatment *T*-scores and the test-retest reliability of the corresponding measure. Although studies commonly report statistical mean differences (e.g., *p* value), standardized mean differences (e.g., Cohen's *d*, Hedge's *g*), and/or measures of relative benefit of one treatment in comparison to another or a control condition (e.g., relative risks, odds ratios), these statistics are often widely misinterpreted and, more importantly, fail to establish a measure of *clinical* significance because they do not answer the question, "What are the chances of a (youth) participant getting better, or worse, after completing (a given) treatment"? It is also important to note that these otherwise commonly reported measures of statistical significance are based on the *mean* performance of a *group*. These statistical measures fail to communicate how an *individual* changes over the course of treatment and how *meaningful* that change is in light of the construct(s) measured. Although the RCI is not a

sufficient substitute for reliable measurement of group performance, it is a critical method to evaluate individual change over time.

Above and below the line of no change ($y = x$; solid line), a set of confidence intervals (dashed lines) were established thereby marking the RCI. More specifically, the area contained within the RCI denoted the range in which changes between pre- and post-treatment mean *T*-scores could vary due to random fluctuation or test error. The RCI values for each subscale were obtained from the Conners-3 manual (Conners, 2008). The manual also explained that changes in scores that meet or exceed the RCI value can be considered to be a statistically significant change 90% of the time (i.e., $p < .10$). In this study, cases that appeared below the RCI indicated a statistically reliable improvement where as those scores that appeared above the RCI indicated a statistically reliable deterioration. Cases that appeared on the line of no change were categorized as "unchanged", and cases that fell between the bounds of the RCI indicated changes that were not reliable and possibly false positives or false negatives.

Research Design

This study is a within-subjects quasi-experimental repeated-measures design.

Analytic Plan

Power Analysis

To determine adequate sample size for the proposed study, an a priori power analysis was conducted using G*Power 3.1 software (Faul et al., 2009). To detect significance of a moderate effect size (dz = 0.5) with a power of 0.80 at an alpha of 0.05, the power analysis indicated a need for 34 participants in order for the paired sample *t*-test analysis to be

adequately powered. To detect significance of a moderate effect size (f = 0.25) with a power of 0.80 at an alpha of 0.05, the power analysis indicated a need for 28 participants in order for the repeated measures within-between ANOVA analyses to be adequately powered. A sensitivity analysis indicated a sample size of 48, with concurrent pharmacotherapy group (ADHD, ADHD + Other, Other, and None) as the between-subjects factor and six measurements across time, would yield an 80% chance of detecting a small-to-moderate effect (f = 0.19). A second sensitivity analysis indicated a sample size of 48, with concurrent pharmacotherapy group as the between-subjects factor and five measurements (DRC score) across time, would yield an 80% chance of detecting a small-to-moderate effect (f = 0.20).

Preliminary Analysis

Prior to conducting the paired samples *t*-test, correlations between pre-treatment parent-reported variables of child functioning were calculated to examine the extent of correlated observations (Pituch et al., 2013). In the absence of any remarkable findings, scores across subscales and between measures were considered to be discrete domains of child functioning and sources of incremental validity. Case analyses were completed to identify potential outliers and influential observations across each remaining outcome variable of interest (i.e., frequency of positive and negative STP behaviors, points earned in the academic learning center, and percentage of DRC target behaviors achieved). Standardized residuals with absolute values larger than 3 indicated outlying values. A sensitivity analysis was then completed to determine the effect of the outlier(s). Tests of main effects and interactions were completed with outlier(s) and without outlier(s). If there was a change in significance in the *p*

values for the main effect(s) or interaction (e.g., significant value with outlier(s) included and nonsignificant value with outlier(s) excluded; or nonsignificant value with outlier(s) included and significant value with outlier(s) excluded), the outlier(s) were noted as having excessive influence and were omitted from the data set prior to completing the primary analyses.

Assumptions: Repeated Measures ANCOVA. A repeated measures ANCOVA was used to examine the differences in mean scores (continuous dependent measure) among multiple groups across time. The repeated measures ANCOVA model also included a covariate (pre-treatment BPM score) to reduce error variability and adjust outcome scores for initial group differences. Prior to conducting the primary analysis, the data was assessed for validity of the repeated measures ANCOVA assumptions: sphericity, homogeneity of regression slopes, linearity, independence, homogeneity of variance, and normality. Mauchly's *W* test was used to determine whether the sphericity assumption had been violated. If the assumption was not met, as indicated by a significant *p* value for Mauchly's *W*, the differences in variances between all levels of the given dependent variable were unequal. This, in turn, would inflate the *F* statistic and rate of type I error. When the Greenhouse-Geisser estimated epsilon (ε) values were ≤ 0.75, the Greenhouse-Geisser degrees of freedom correction was applied to the corresponding repeated measures test. If the Greenhouse-Geisser estimated ε values were > 0.75, however, the Huynh-Feldt degrees of freedom correction was applied as this correction has been reported to be more efficient and powerful than the Greenhouse-Geisser correction (Abdi, 2010). In the absence of a significant interaction effect, the homogeneity of regression slopes assumption was met. Scatter plots with regression lines for different groups indicated similar

regression slopes across groups. In the absence of significant deviations from linearity in the regression lines, the linearity assumption was met. The study design was also reviewed with particular attention to the ways in which children could obtain points across the camp day. There was no evidence to suggest a violation of the independence assumption. Levene's test, which can be applied to a set of data points that lack normality and still retain strong power (Gastwirth, Gel, & Miao, 2009), was then used to determine if there was a significant difference in the residual variances across groups. If the results of the Levene's test indicated that the homogeneity of variance assumption had been violated, a log transformation was applied to the data. In cases were a dependent variable had a value of '0', '1' was added to each score before the log transformation was applied. If Levene's test still indicated homogeneity of variance had been violated, outliers with z scores greater than 3 were omitted. Despite the persistent violation of the homogeneity of variance assumption, the repeated measures ANCOVA test was still conducted, as this statistical test has been shown to be robust against unequal variances provided there are closet to equal group sizes. The distributions of the residuals for each group indicated lack of normality, once again due to skewness. However, ANCOVA is robust to the violation of normality.

Primary Analysis

Univariate and multivariate tests were then conducted to investigate the proposed research questions. A paired samples *t*-test was used to examine the difference in mean scores of pre-treatment and post-treatment parent-reported measures of child functioning. The significance of these differences were evaluated at the $p = .05$ level. Repeated measures

ANCOVA models, absent the nonsignificant interaction term, yielded *F* values which were then examined for significance at the $p = .05$ level. If the *p* value of an *F* test met this level of significance, post-hoc tests were used to examine specific contrasts of interest. Adjusted means and their standard errors were examined to determine which group(s) differed on outcome scores, and *p* values from the pairwise comparisons table were used to determine which group differences were statistically significant. To protect against family-wide type I error inflation due to the number of pairwise comparisons, a Bonferroni-Holm procedure was conducted for each repeated measures ANCOVA (Holm, 1979). The post-hoc multivariate analyses provided four test statistics associated with the *F* ratio. Pillai's Trace statistics is considered to be the most powerful and robust of these four (Carey, 1998), and thus, Pillai's Trace was reported for the corresponding analyses.

Chapter IV: Results

The following statistical analyses were computed to: (1) investigate the effectiveness of the 6-week adaptation of the STP model by examining behavioral improvement, as defined by (a) differences in pre- and post-treatment parent-reported measures of child inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations; (b) change in frequency of behaviors across standard STP point system categories, and (c) change in percentage of daily report card target behaviors achieved, and (2) to investigate whether children who receive concurrent pharmacotherapy experience greater benefits from the 6-week STP. Statistical analyses were computed using IBM SPSS Statistics (Version 28) predictive analytics software.

Missing Data

Parent-reported measures of baseline child functioning were missing in 12 cases for the Conners-3-P(S) rating scale and in 4 cases for the BPM rating scale. Values for *answering attention questions* could not be generated in 20 cases. As such, these cases were omitted from the data set used in the corresponding repeated measures ANCOVA analysis.

Descriptive Statistics

Descriptive statistics for baseline demographic and clinical characteristics are summarized across Tables 1 and 3. Pre-treatment differences across groups were examined by means of one-way ANOVAs (continuous variables) and chi-square analysis (categorical data) to assess for any significant differences in baseline demographic and clinical characteristics of participants across groups, which may otherwise confound subsequent findings. Pre-treatment

comparisons between the Conners-3-P(S) subscales and the BPM were all nonsignificant, indicating similar severity scores on each of the parent-reported measures of child functioning at baseline.

Relationship Between Dependent Measures

Correlations between total family income and all pre-treatment parent-reported variables of child functioning were calculated and reviewed to better understand the relationship between the outcome variables of interest (Table 4). Correlations were largest between the executive functioning and inattention subscales of the Conners-3-(P)S ($r = .688$; $p < .001$) indicating a 6.89% overlap in shared variance between these two variables Correlations between the executive functioning and hyperactivity/impulsivity subscales ($r = .427$, $p = .006$) and the inattention and hyperactivity/impulsivity subscales ($r = .422$; $p = .007$) indicated medium to large effect sizes with 4.27% overlap in shared variance between the executive functioning and hyperactivity/impulsivity subscales and a 4.22%. overlap in shared variance between the inattention and hyperactivity/impulsivity subscales. The statistically significant relationship between these variables of child functioning suggest a fairly predictable relationship between the constructs, as expected.

There was also a statistically significant correlation between the defiance/aggression and peer relations subscales ($r = .360$; $p = .022$) and a 3.6% overlap in shared variance between these two variables. In accordance with the literature, this correlation suggests a fairly predictable relationship between defiance/aggression and the way in which those functional impairments, in turn, negatively affect the ability to establish and maintain peer relationships.

Analyses also revealed a statistically significant negative correlation between the defiance/aggression and hyperactivity/impulsivity subscales ($r = -.387$; $p = .014$). Although children with ADHD may demonstrate occasional irritability, anger, or dangerous behaviors, presenting concerns of persistent patterns of defiance and/or aggression are often better characterized by the Oppositional Defiant Disorder (ODD) diagnosis (APA, 2013). The negative correlation observed within this study sample suggests children generally demonstrated greater impairment in behaviors related to hyperactivity/impulsivity (characteristic of ADHD) than those behaviors related to defiance/aggression (characteristic of ODD).

Lastly, there was a statistically significant correlation between total family income and learning problem scores at pre-treatment ($r = .319$; $p = .045$). This, in conjunction with the sub-threshold scores within this domain of functioning at both pre- and post-treatment, provides further support for the existing literature on socioeconomic status as a protective factor for academic outcomes. Correlation values between the remaining parent-reported outcome variables were nonsignificant, indicating that each variable provided incremental validity.

Table 3

Descriptive Statistics: Pre-Treatment Parent-Reported Measures of Child Functioning

	Total Sample (N = 40)	ADHD Only (*n* = 18)	ADHD + Other (*n* = 5)	Other (*n* = 9)	None (*n* = 8)	*F*	*p*
Conners-3-P(S) (Mean, *SD*)							
Inattention	75.13, *11.98*	74.50, *10.23*	71.4, *14.15*	76.00, *16.45*	77.88, *10.05*	.317[1,2]	.813
Hyperactivity/Impulsivity	76.25, *11.83*	75.56, *13.33*	73.60, *6.80*	76.33, *13.59*	79.38, *9.64*	.274[1,2]	.843
Learning Problems	57.53, *12.77*	60.06, *13.71*	48.60, *6.99*	59.67, *10.32*	55.00, *14.63*	1.267[1,2]	.300
Executive Functioning	68.85, *12.42*	67.78, *12.37*	69.20, *10.77*	71.89, *14.59*	67.63, *12.72*	.237[1,2]	.870
Defiance/Aggression	63.28, *15.23*	62.72, *14.42*	65.80, *10.64*	62.44, *18.36*	63.88, *18.21*	.062[1,2]	.980
Peer Relations	73.90, *15.70*	72.61, *15.48*	84.00, *9.59*	74.44, *18.68*	69.88, *15.62*	.903[1,2]	.449
BPM (Mean, *SD*)	17.23, *5.66*	17.05, *4.92*	16.63, *9.33*	18.82, *4.35*	16.13, *5.06*	.412[1,3]	.745

Note. [1]$df_b = 3$; [2]$df_w = 39$; [3]$df_w = 44$. ADHD Only = prescribed FDA-approved medication for ADHD; ADHD + Other = prescribed FDA-approved medication for ADHD in addition to one or more psychotropics, with indication(s) other than ADHD; Other = prescribed one or more psychotropics, with indication(s) other than ADHD; None = no medications prescribed.

Table 4

Means, Standard Deviations, and Correlations of Total Family Income and Pre-Treatment Continuous Variables

	M	*SD*	1	2	3	4	5	6	7	8
1. Total Family Income	239,647.10	216,033.79	--							
2. Inattention	75.12	11.98	.172	--						
3. Hyperactivity/Impulsivity	76.25	11.83	.206	.422**	--					
4. Learning Problems	57.52	12.76	.319*	.287	-.046	--				
5. Executive Functioning	68.85	12.42	.099	.688**	.427**	.242	--			
6. Defiance/Aggression	63.27	15.23	-.023	-.387*	.199	-.163	-.228	--		
7. Peer Relations	73.90	15.70	.142	-.122	.159	-.020	.180	.360*	--	
8. BPM	17.23	5.66	-.082	.170	-.068	-.019	.198	-.007	-.065	--

Note. M and *SD* are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.

Table 5

T-Score Guidelines: Conners-3-P(S)

T-score	Percentile	Guideline
70+	98 +	Very elevated score *(many more concerns than are typically reported)*
60 – 69	84 – 97	Elevated score *(more concerns than are typically reported)*
40 – 59	16 – 83	Average score *(typical levels of concern)*
< 40	< 16	Low Score (*fewer concerns than are typically reported)*

Primary Analyses: Tests of Research Questions

Research Question 1

Do children who complete the 6-week STP show improvements in parent-reported outcomes, including: inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations?

Hypothesis 1. Parent-reported scores of inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations will significantly decrease following completion of the 6-week STP.

Results. A paired samples *t*-test was conducted to determine whether mean *T*-scores on Conners-3-P(S) subscales at pre-treatment differed from mean scores reported at post-treatment (Table 6). This analysis included 40 participants. There was a significant mean difference between pre- and post-treatment *T*-scores on the inattention ($t = 3.36$, $p = .002$; Cohen's $d = .53$) and hyperactivity/impulsivity ($t = 2.56$, $p = .014$; Cohen's $d = .41$) subscales, each demonstrating a minimum effect size (Ferguson, 2009). There was also a significant difference in mean *T*-scores on the peer relations ($t = 2.15$, $p = .038$; Cohen's $d = .34$) subscale, however these differences did not achieve the level of recommended minimum effect size to indicate a practically significant effect. At post-treatment, mean *T*-scores on the inattention, hyperactivity/impulsivity, and peer relations subscales remained in the 'clinical' range ($T \geq 65$).

A one-way ANOVA was conducted to compare post-treatment mean *T*-scores across concurrent pharmacotherapy groups (Table 7). There was a significant difference in learning

problems scores (p = .022; η^2 = .232) demonstrating a large effect size. Post-hoc multiple comparisons indicated a significant difference in mean learning problems *T*-scores between the ADHD Only and ADHD + Other group (p = .043); ADHD Only and None group (p = .018); ADHD + Other and Other group (p = .036); and the Other and None group (p = .018) (Figure 1). To further investigate the clinical significance of these findings, pre- and post-treatment learning problems scores were compared by referencing the range of *T*-scores designated as 'clinical' and 'subclinical' per the Conners-3-P(S) manual. The ADHD Only group remained in the low end of the subclinical range. On average, children in the ADHD + Other and Other groups demonstrated a slight increase in parent-reported scores of learning problems at post-treatment. Despite the slight increase, the ADHD + Other group remained below the subclinical threshold. However, the increase was of clinical relevance for the Other group in that mean learning problem scores at post-treatment crossed the subclinical range threshold. On average, participants in the ADHD Only (M = 60.00, SD = 9.95) and Other (M = 61.44, SD = 11.71) groups scored in the subclinical range. Participants in the ADHD + Other (M = 50.00, SD = 6.89) and None (M = 50.00, SD = 6.89) groups scored below the threshold of subclinical *T*-scores, indicating a post-treatment level of functioning closer to the mean of the normal sample.

Table 6

Paired Samples T-Test Analysis for Conners-3-P(S) Subscale T-Scores at Pre-and Post-Treatment

	Pre-Treatment		Post-Treatment		*r*	*t*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*				
Inattention	75.13	11.98	70.25	12.79	.727	3.36	.002*	.53
Hyperactivity/Impulsivity	76.25	11.83	71.28	13.80	.551	2.56	.014*	.41
Learning Problems	57.53	12.76	57.10	10.32	.730	0.31	.761	.05
Executive Functioning	68.85	12.42	67.10	12.69	.593	0.98	.335	.15
Defiance/Aggression	63.28	15.23	63.70	16.79	.463	-0.16	.873	-.03
Peer Relations	73.90	15.70	69.60	18.08	.728	2.15	.038*	.34

Note. Clinical scores: *T*-score ≥ 65. Subclinical scores: *T*-score = 60 – 64. Cohen's *d* is an effect size estimate of group difference; values of .41 indicate the recommended minimum effect size representing a "practically" significant effect for social science data; values of 1.15 indicate a moderate effect; and values of 2.70 indicate a strong effect (Ferguson, 2009).

Table 7

Descriptive Statistics: Post-Treatment Parent-Reported Measures of Child Functioning

	Total Sample (N = 40)	ADHD only (*n* = 18)	ADHD + Other (*n* = 5)	Other (*n* = 9)	None (*n* = 8)	*p*
Conners-3-P(S) *T*-Score (Mean, *SD*)						
Inattention	70.25, *12.79*	70.39, *10.87*	64.40, *15.89*	71.78, *14.85*	71.88, *14.05*	.743
Hyperactivity/Impulsivity	71.28, *13.80*	74.22, *12.31*	71.40, *16.58*	66.89, *14.73*	69.50, *15.41*	.618
Learning Problems	57.10, *10.32*	60.00, *9.95*	50.00, *6.89*	61.44, *11.71*	50.13, *5.64*	.022*
Executive Functioning	67.10, *12.69*	68.83, *12.86*	62.40, *12.64*	71.56, *10.81*	61.13, *13.45*	.278
Defiance/Aggression	63.70, *16.79*	61.67, *14.90*	70.60, *20.84*	69.67, *18.06*	57.25, *16.50*	.340
Peer Relations	69.60, *18.08*	72.06, *12.24*	75.80, *18.46*	63.67, *20.89*	66.88, *17.62*	.574

Figure 1

Bar Graph of Post-Treatment T-Scores: Conners-3-P(S) Learning Problems Subscale

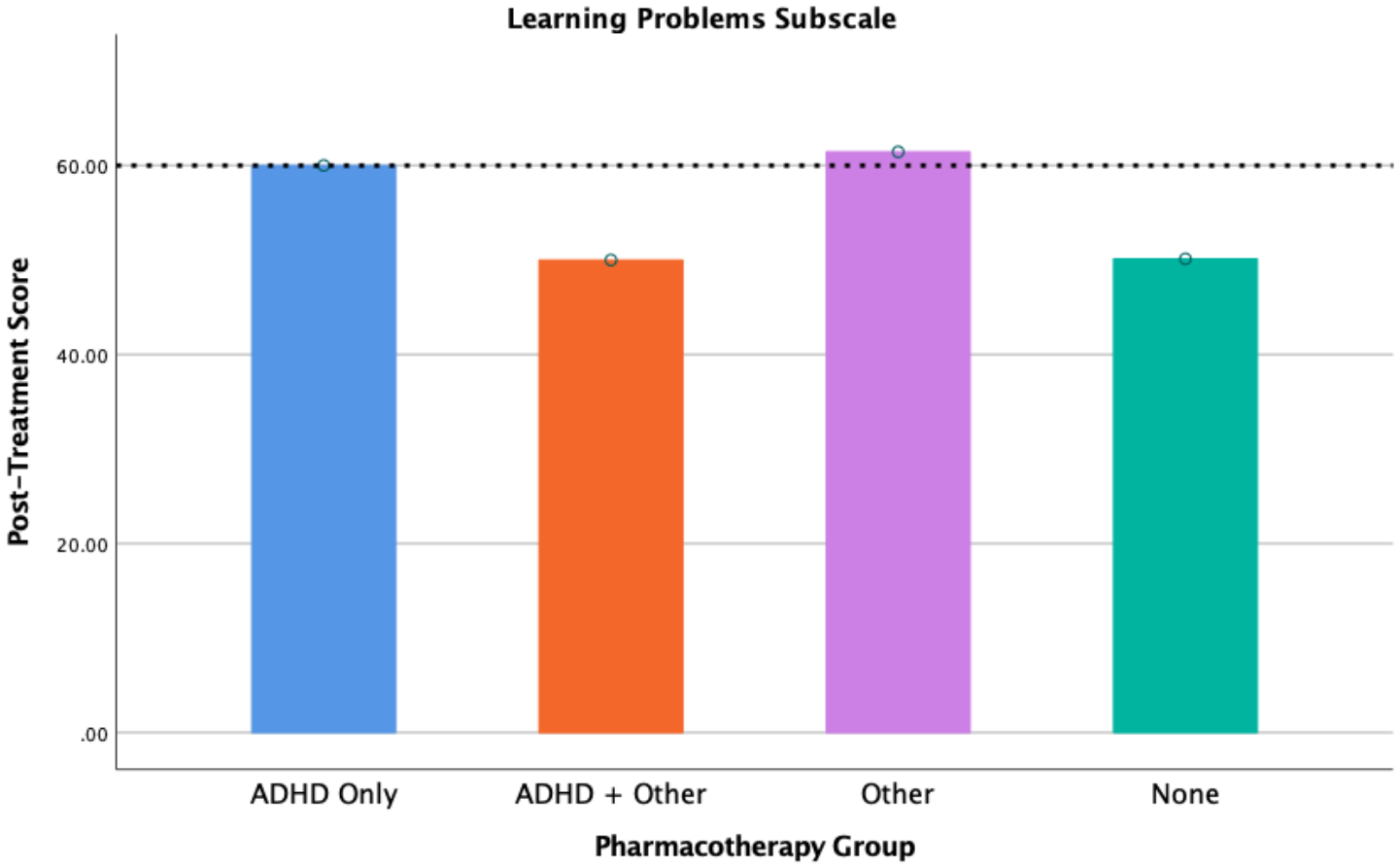


Table 8

Reliability Change Index: Status at Post-Treatment by Concurrent Pharmacotherapy Group

	Total Sample (N = 40)					ADHD Only (*n* = 18)					ADHD + Other (*n* = 5)					Other (*n* = 9)					None (*n* = 8)				
	Reliable +	*Improvement*	*Unchanged*	*Deterioration*	*Reliable –*	*Reliable +*	*Improvement*	*Unchanged*	*Deterioration*	*Reliable –*	*Reliable +*	*Improvement*	*Unchanged*	*Deterioration*	*Reliable –*	*Reliable +*	*Improvement*	*Unchanged*	*Deterioration*	*Reliable –*	*Reliable +*	*Improvement*	*Unchanged*	*Deterioration*	*Reliable –*
1	14	11	3	9	3	4	8	2	4	–	3	–	–	1	1	4	1	–	3	1	3	2	1	1	1
	35.0% [a]			7.50% [b]		22.2% [a]			0% [b]		60.0% [a]			20.0% [b]		44.4% [a]			11.1% [b]		37.5% [a]			12.5% [b]	
2	18	6	4	5	7	5	5	2	2	4	2	–	–	–	3	6	–	1	2	–	5	1	1	1	–
	45.0% [a]			17.5% [b]		27.8% [a]			22.2% [b]		40.0% [a]			60.0% [b]		66.7% [a]			0% [b]		62.5% [a]			0% [b]	
3	4	15	11	7	3	2	6	6	1	3	–	4	–	1	–	1	4	–	4	–	1	1	5	1	–
	10.0% [a]			7.50% [b]		11.1% [a]			16.7% [b]		0% [a]			0% [b]		11.1% [a]			0% [b]		12.5% [a]			0% [b]	
4	6	15	4	10	5	1	7	2	5	3	2	1	–	2	–	1	4	–	2	2	2	3	2	1	–
	15.0% [a]			12.5% [b]		5.6% [a]			16.7% [b]		40.0% [a]			0% [b]		11.1% [a]			22.2% [b]		25.0% [a]			0% [b]	
5	15	3	7	2	13	6	1	4	–	7	2	1	–	1	1	6	1	–	1	1	1	–	3	–	4
	37.5% [a]			32.5% [b]		33.3% [a]			38.9% [b]		40.0% [a]			20.0% [b]		66.7% [a]			11.1% [b]		12.5% [a]			50.0% [b]	
6	10	9	11	6	4	3	3	6	4	2	2	–	2	–	1	3	4	1	1	–	2	2	2	1	1
	25.0% [a]			10.0% [b]		16.7% [a]			11.1% [b]		40.0% [a]			20.0% [b]		33.3% [a]			0% [b]		25.0% [a]			12.5% [b]	

Note. 1 = Inattention subscale; 2 = Hyperactivity/Impulsivity subscale; 3 = Learning Problems subscale; 4 = Executive Functioning subscale, 5 = Defiance/Aggression subscale; 6 = Peer Relations subscale. [a] % Reliably improved. [b] % Reliably deteriorated.

The RC analysis (n = 40) allowed for a statistically sound measure of the individual degree of change, whereby participants were categorized into one of five change categories: *reliable improvement* (score below the lower bound of the RCI), *improvement* (scores below the line of no change), *no change* (scores on the line of no change), *deterioration* (scores above the line of no change), and *reliable deterioration* (scores above the upper bound of the RCI) (Figures 2-7). Percentage rates of reliable improvement and reliable deterioration were calculated for each subscale, across both the total sample and the discrete concurrent pharmacotherapy groups (Table 8). As one sample, the greatest rate of reliable improvement was demonstrated in the hyperactivity/impulsivity domain (45.0%; n = 18) and the greatest rate of reliable deterioration was demonstrated in the defiance/aggression domain (32.5%, n = 13). Within the ADHD Only group, the greatest rates of reliable improvement (33.3%, n = 6) and reliable deterioration appeared within the defiance/aggression domain. Among the ADHD + Other subsample, the greatest rate of reliable improvement was demonstrated in the inattention domain (60.0%; n = 3) and the greatest rate of reliable deterioration was demonstrated in the hyperactivity/impulsivity domain (60.0%; n = 3). Among the participants in the Other group, the greatest rates of reliable improvement occurred in both the hyperactivity/impulsivity and defiance/aggression domains (66.7%; n = 6) while the greatest rates of reliable deterioration occurred in the executive functioning domain (22.2%, n = 2). Participants that did not receive any concurrent pharmacotherapy (None) demonstrated the greatest rate of reliable improvement in the hyperactivity/impulsivity domain (62.5%, n = 5) and the greatest rate of reliable deterioration in the defiance/aggression domain (50.0%, n = 4).

Chi-square analyses were conducted to determine whether rates of reliable improvement on a given Conners-3-P(S) subscale differed from rates of reliable deterioration (Table 9). Within the total sample, there was a statistically significantly greater rate of reliable improvement, compared to reliable deterioration, on the hyperactivity/impulsivity ($x^2 = 6.94$, $p = .008$; V = .417) and defiance/aggression ($x^2 = 11.56$, $p < .001$, V = .537) subscales. The effect size for the hyperactivity/impulsivity rate of reliable improvement met the guidelines for minimum effect size, representing a "practically" significant effect; and the defiance/aggression rate of reliable improvement demonstrated a moderate effect. Within each of these two domains of functioning, one concurrent pharmacotherapy group demonstrated statistically significantly rates of deterioration: the ADHD + Other group demonstrated statistically significantly greater rates of reliable deterioration (60.0%, $n = 3$; $x^2 = 5.00$, $p = .025$; V = 1.000) within the hyperactivity/impulsivity domain, and the ADHD Only group demonstrated statistically significant greater rates of reliable deterioration (38.7%, $n = 7$; $x^2 = 5.73$, $p = .017$; V = .564) within the defiance/aggression domain. Effect sizes for these rates of reliable change ranged from strong (hyperactivity/impulsivity) to moderate (defiance/aggression). The remaining subscales demonstrated nonsignificant rates of reliable change, both as a total sample and across concurrent pharmacotherapy groups.

Table 9

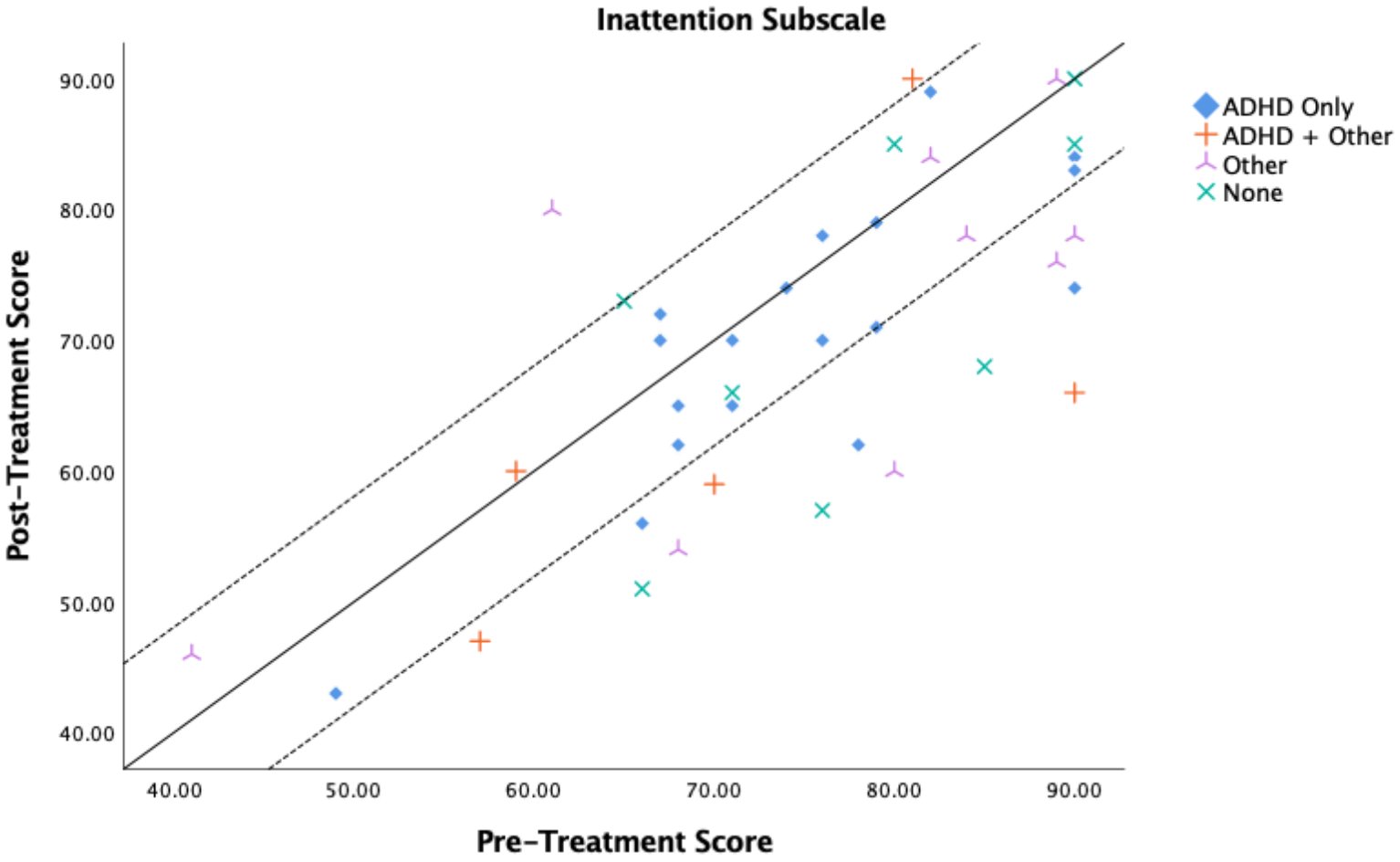
Chi-Square Analysis: Rates of Reliable Change (Improvement / Deterioration)

	Total Sample			ADHD Only			ADHD + Other			Other			None		
	x^2	*p*	Cramer's V	x^2	*p*	Cramer's V	x^2	*p*	Cramer's V	x^2	*p*	Cramer's V	x^2	*p*	Cramer's V
1	1.75	.186	.209	–	–	–	1.88	.171	.612	0.90	.343	.316	0.69	.408	.293
2	6.94	.008**	.417	1.98	.160	.331	5.00	.025*	1.000	–	–	–	–	–	–
3	0.36	.548	.095	0.45	.502	.158	–	–	–	–	–	–	–	–	–
4	1.01	.315	1.59	0.21	.645	.108	–	–	–	0.32	.571	.189	–	–	–
5	11.56	< .001**	.537	5.73	.017*	.564	0.83	.361	.408	2.25	.134	.500	1.14	.285	.378
6	1.48	.224	.192	0.45	.502	.158	0.83	.361	.408	–	–	–	0.38	.537	.218

Note. 1 = Inattention subscale; 2 = Hyperactivity/Impulsivity subscale; 3 = Learning Problems subscale; 4 = Executive Functioning subscale, 5 = Defiance/Aggression subscale; 6 = Peer Relations subscale. Cramer's V is an effect size estimate of strength of association; values of .2 indicate the recommended minimum effect size representing a "practically" significant effect for social science data; values of .5 indicate a moderate effect; and values of .8 indicate a strong effect (Ferguson, 2009). * indicates $p < .05$. ** indicates $p < .01$.

Figure 2

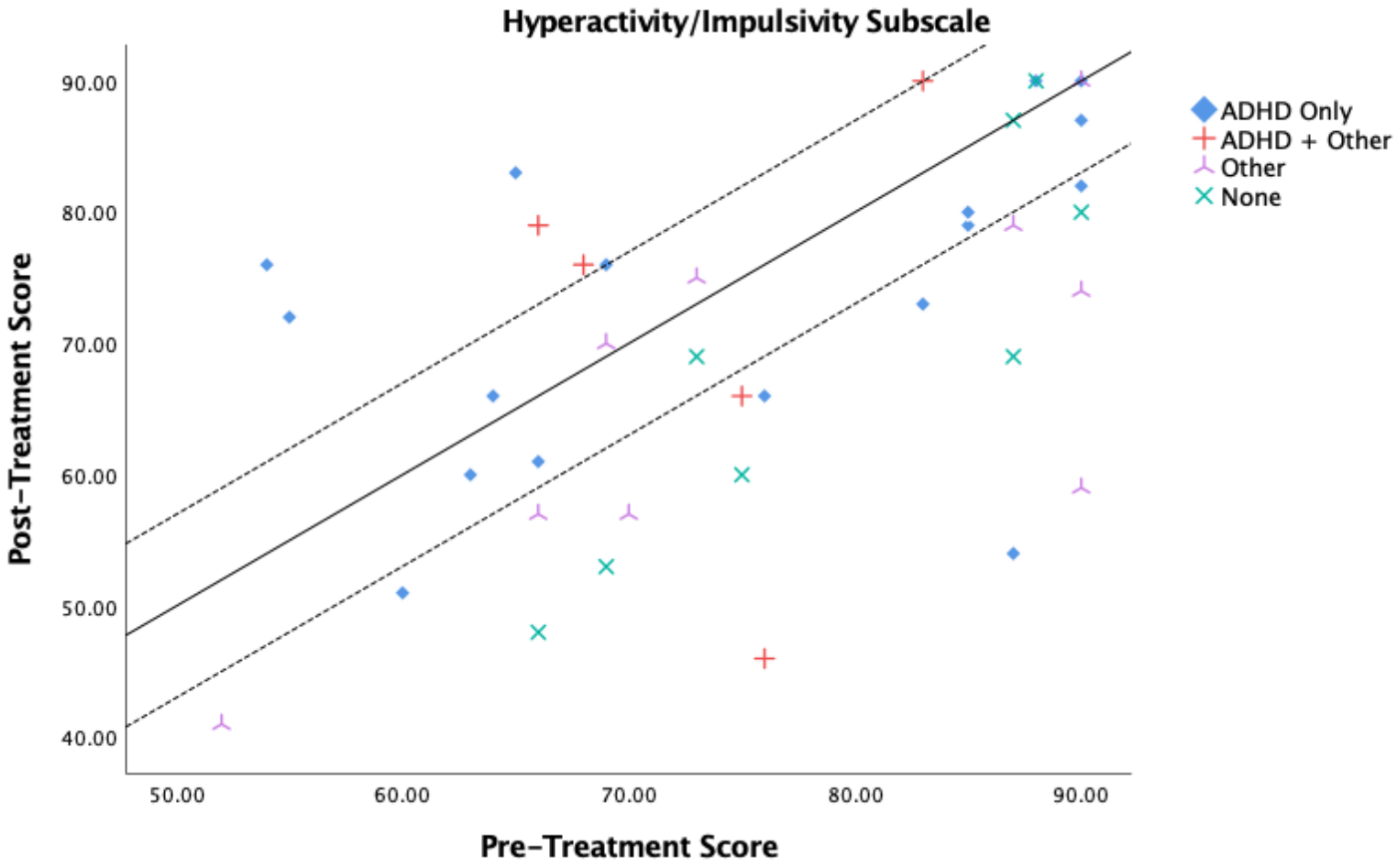
Scatter Plot of Pre- and Post-Treatment T-Scores: Conners-3P(S) Inattention Subscale



Note. Solid diagonal line indicates line of no change. The dashed diagonal lines indicate the upper and lower bounds of the Reliable Change Index (RCI). RCI (Inattention *T*-score) = 8.06.

Figure 3

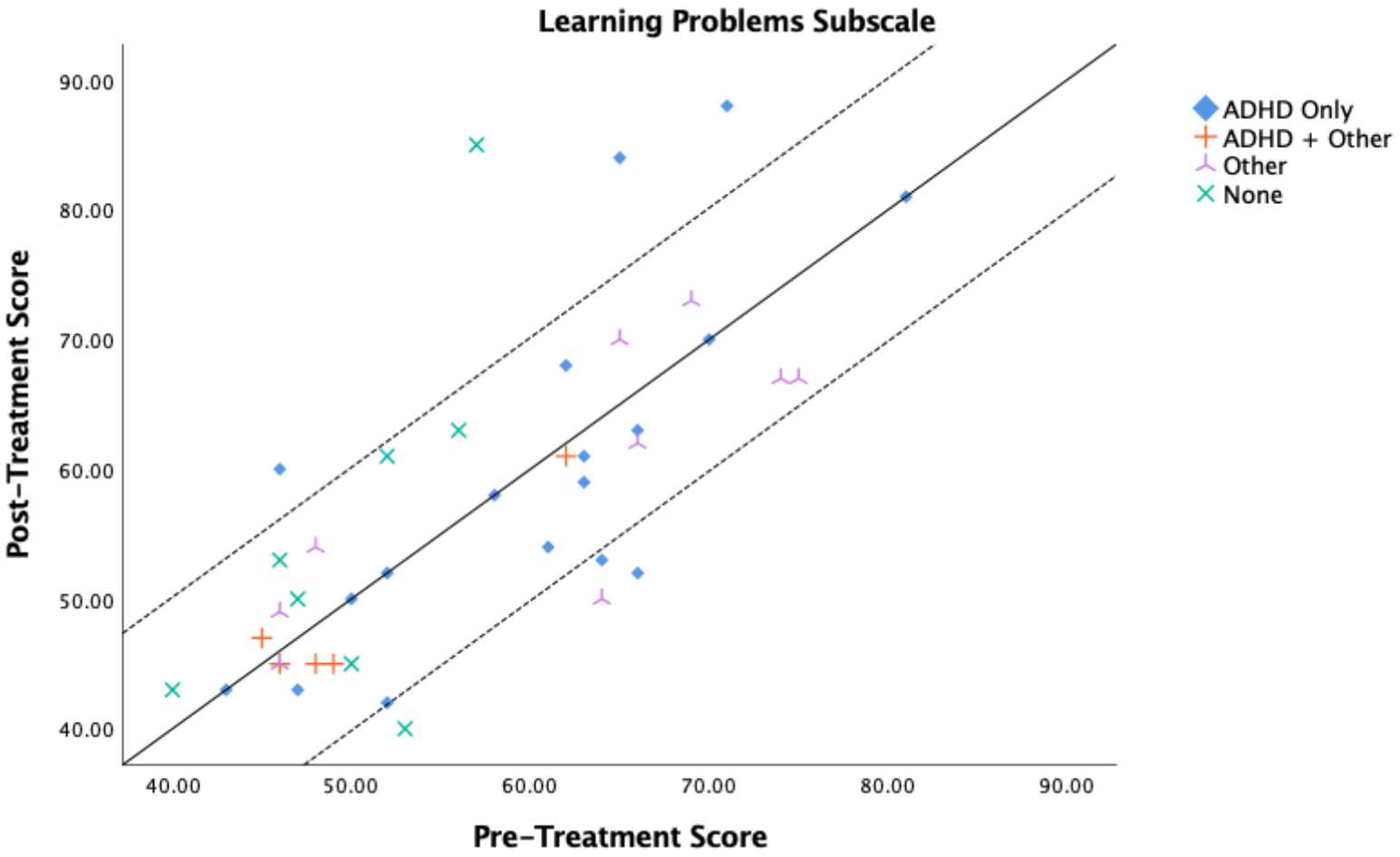
Scatter Plot of Pre- and Post-Treatment T-Scores: Conners-3P(S) Hyperactivity/Impulsivity Subscale



Note. RCI (Hyperactivity/Impulsivity *T*-score) = 6.98.

Figure 4

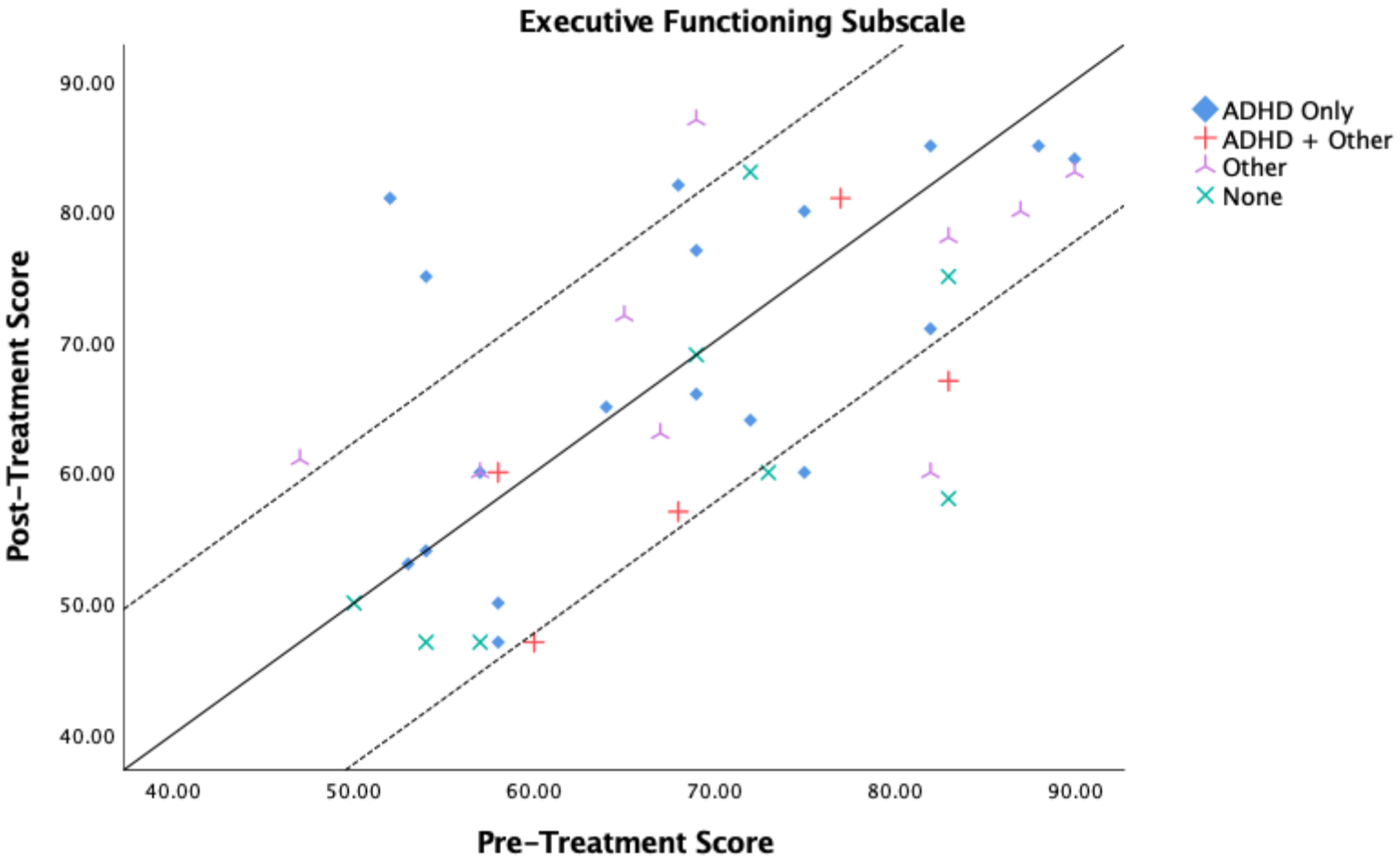
Scatter Plot of Pre- and Post-Treatment T-Scores: Conners-3P(S) Learning Problems Subscale



Note. RCI (Learning Problems *T*-score) = 10.14.

Figure 5

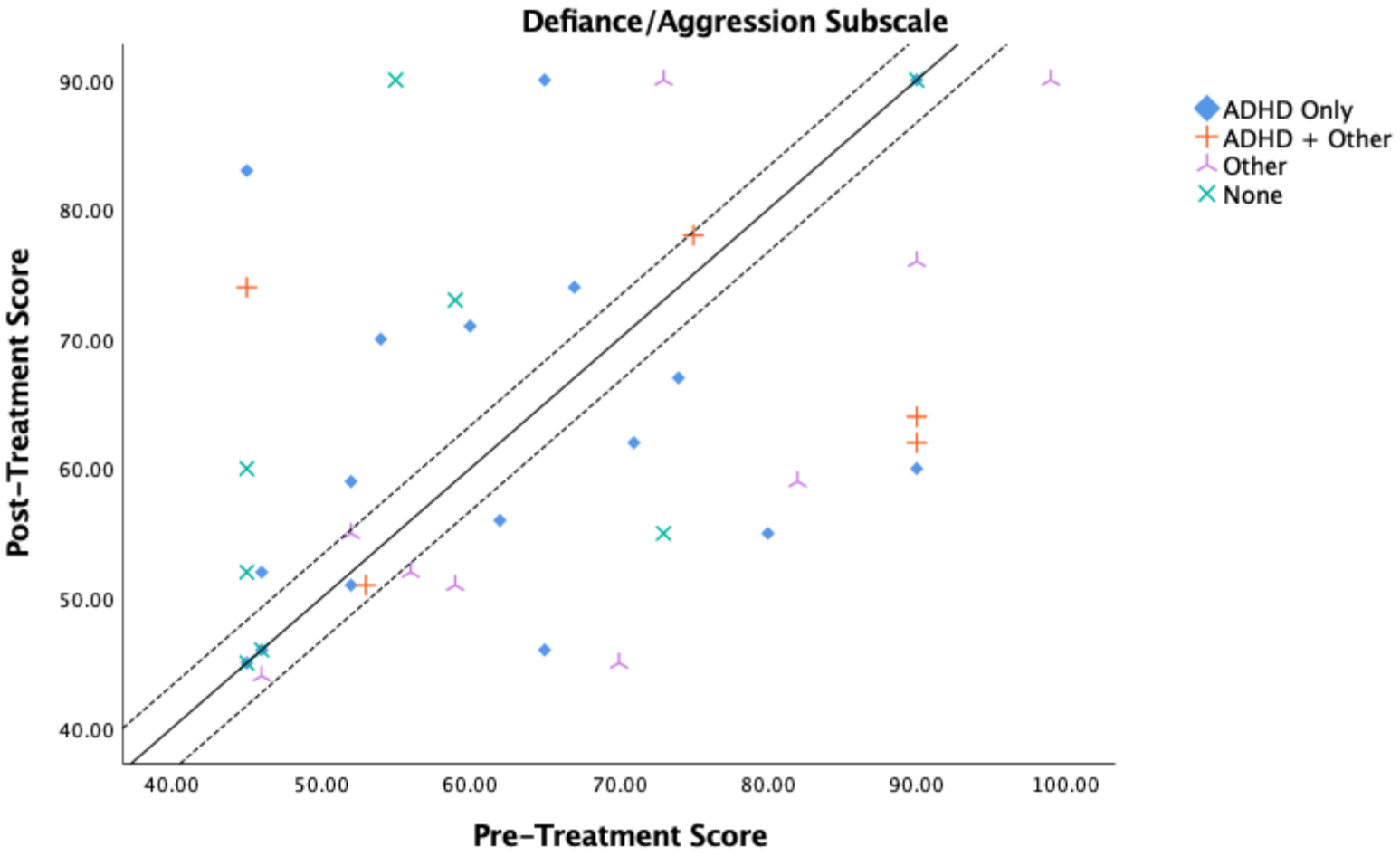
Scatter Plot of Pre- and Post-Treatment T-Scores: Conners-3P(S) Executive Functioning Subscale



Note. RCI (Executive Functioning *T*- Score) = 12.31.

Figure 6

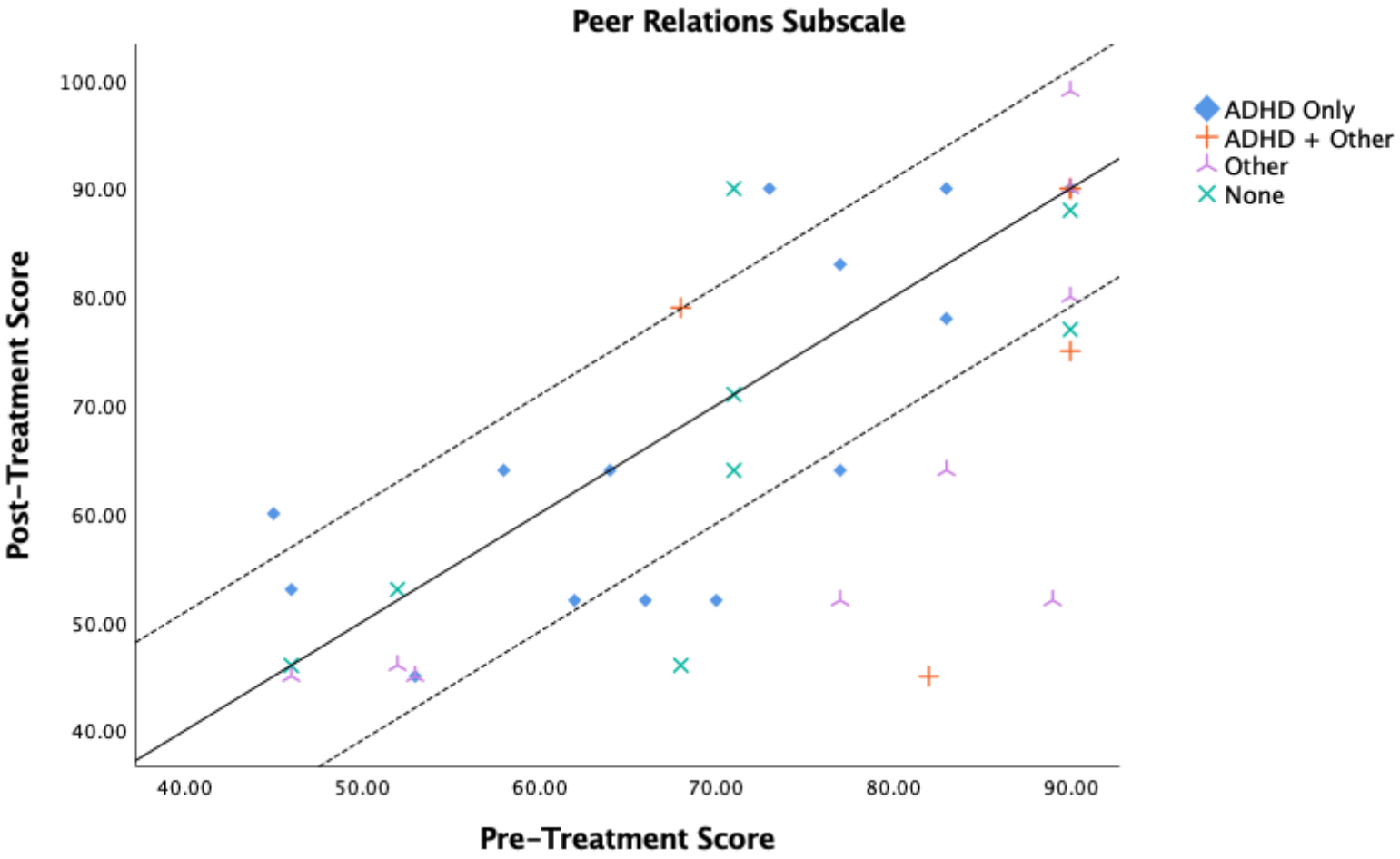
Scatter Plot of Pre- and Post-Treatment T-Scores: Conners-3P(S) Defiance/Aggression Subscale



Note. RCI (Defiance/Aggression *T*-score) = 3.29.

Figure 7

Scatter Plot of Pre- and Post-Treatment T-Scores: Conners-3P(S) Peer Relations Subscale



Note. RCI (Peer Relations *T*-score) = 10.91.

Summary of Results. Compared to the mean *T*-scores reported at the outset of treatment, there was a statistically significant decrease in mean *T*-scores at post-treatment on the inattention (-4.88 points; $d = .53$), hyperactivity/impulsivity (-4.97 points; $d = .41$) and peer relations (-4.30 points; $d = .34$) subscales. Of these significant differences, only the decreases in mean *T*-scores on the inattention and hyperactivity/impulsivity subscales demonstrated a minimum effect size. At post-treatment, concurrent pharmacotherapy groups demonstrated statistically significant differences in mean *T*-scores on the learning problems subscale. Specifically, the ADHD + Other group demonstrated a significantly lower score than both the ADHD Only ($p = .043$) and Other groups ($p = .036$); and the None group also demonstrated a significantly lower score than both the ADHD Only ($p = .018$) and Other groups ($p = .018$). The RC analysis allowed for a statistically sound measure of individual participants' degree of change subsequent to the 6-week STP. Within the total sample, rates of reliable improvement ranged from 45.0% ($n = 18$; hyperactivity/impulsivity) to 10.0% ($n = 4$; executive functioning); rates of reliable deterioration ranged from 32.5% ($n = 13$; defiance/aggression) to 7.5% ($n = 3$; inattention and learning problems); and rates of no change ranged from 27.5% ($n = 11$; learning problems) to 7.5% ($n = 3$; inattention).

Research Question 2

Do children who complete the 6-week STP show significant improvements in frequency of positive behaviors, including: answering attention questions, following activity rules, complying with counselor commands, contributing to group discussions, ignoring a negative stimulus, helping a peer, and sharing with a peer?

Hypothesis 2. The frequency of behaviors within each positive behavior category will significantly increase throughout the 6-week STP.

Results. To test this hypothesis, a repeated measures ANCOVA model with concurrent pharmacotherapy group as the between-subjects factor, time as the within-subjects factor, and the pre-treatment BPM total score as the covariate was used. Separate repeated measures ANCOVA models were used to examine change in frequency counts across each of the seven positive behavior categories on the STP point system. Case analyses revealed outlying values, within the *complying with counselor commands*, *ignoring a negative stimulus*, *helping a peer*, and *sharing with a peer* positive behavior categories (Table 10).

Table 10

Case Analyses: Standardized Residuals Across Positive Behavior Categories

ID	BPM Score	Group	Standardized Residual					
			Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Complying with Command								
36	17	None		4.02				
Ignoring Negative Stimulus								
19	8	ADHD + Other			3.31			
Helping a Peer								
49	17	ADHD Only					4.74	3.16
Sharing with a Peer								
6	13	Other				3.20		
40	20	ADHD Only		3.00	3.80	3.72		
49	17	ADHD Only					5.25	

A sensitivity analysis was completed to determine the impact of outlying values on the ANCOVA tests results (Table 11). In the absence of the outlier in the *helping a peer* group, the

ANCOVA yielded a significant main effect (p = .006) and interaction term (p = .026), whereas both main effects and the interaction term were nonsignificant when the outlier was included. Similarly, when outliers were omitted from the *sharing with a peer* group, the ANCOVA yielded a significant interaction term (p < .001) that was nonsignificant when the outliers were included. These discrepancies indicated that the outlying cases in the *helping a peer* and *sharing with a peer* behavior categories demonstrated excessive influence on the results and were thus omitted from the data set used to run the primary ANCOVA analysis. The analysis included 45 participants, with the exception of the *answering attention questions* sample (n = 15).

Table 11

Sensitivity Analyses: Main Effects and Interaction

	Tests of Between-Subjects Effects		Excessive
	Included	Excluded	Influence?
Source	*p*	*p*	
Complying with Command			No
Group	.018*	.018*	
BPM	.397	.397	
Group*BPM	.032*	.032*	
Ignoring a Negative Stimulus			No
Group	.064	.167	
BPM	.281	.354	
Group*BPM	.123	.207	
Helping a Peer			Yes
Group	.074	.006*	
BPM	.572	.463	
Group*BPM	.129	.026*	
Sharing with a Peer			Yes
Group	.030*	<.001**	
BPM	.686	.943	
Group*BPM	.031	<.001**	

Note. *p < 0.05; **p < .001

Levene's test was used to determine homogeneity of variances for mean frequency counts within each of the seven positive behavior categories in the STP point system. The results indicated that the error variance of mean frequency counts in each of the positive behavior categories was similar across the weeks of treatment, with the exception of three categories: *complying with counselor command* (p = .011; p =.022), *contributing to group discussion* (p = .039), and *helping a peer* (p = .020; p = .002; p = .007) (Table 12). To address the violation in the homogeneity of variance assumption across these three positive behavior categories, a log transformation was applied. In cases where a dependent variable had a value of '0', a '1' was added to the score before the log transformation was applied.

Table 12

Levene's Test: Positive Behavior Categories

Behavior Category		F	df1	df2	Sig.
Complying with Counselor Command					
	Week 1	4.243	3	40	.011*
	Week 3	3.578	3	40	.022*
Contributing to Group Discussion					
	Week 6	3.054	3	41	.039*
Helping a Peer					
	Week 1	3.680	3	41	.020*
	Week 4	5.867	3	41	.002*
	Week 5	4.641	3	41	.007*

The primary repeated measures ANCOVA was then conducted. Mauchly's *W* was significant for five behavior categories, indicating that the assumption of sphericity had been

violated (Table 13). Based on the estimated sphericity (ε), the degrees of freedom was corrected using either the Greenhouse-Geisser or Huynh-Feldt correction.

Table 13

Mauchly's Test: Positive Behavior Categories

Behavior Category	Mauchly's *W*	*p*	Epsilon (ε)	
			Greenhouse-Geisser	Huynh-Feldt
Complying with Counselor Command[1]	.088	<.001	.610	.797
Contributing to Group Discussion[2]	.361	.001	.776	1.00
Ignoring a Negative Stimulus[1]	.276	<.001	.687	.909
Helping a Peer[1]	.281	<.001	.671	.886
Sharing with a Peer[1]	.116	<.001	.500	.641

Note. [1]Greenhouse-Geisser correction indicated. [2]Huynh-Feldt correction indicated.

The results of the Tests of Within-Subjects Effects revealed a significant main effect of time for *answering attention questions* (F (5, 50) = 2.69; p = .031), *following activity rules* (F (3.50, 129.65) = 5.61; p < .001), *contributing to group discussion* (F (5, 143.52) = 2.95; p = .014); and *helping a peer* (F (3.36, 124.16) = 2.86; p = .034) behavior categories (Table 14). The significant differences in mean frequency counts across the *following activity rules* (η^2 = .132), *contributing to group discussion* (η^2 = .074), and *helping a peer* (η^2 = .072) categories indicated medium effect sizes; and the significant differences in mean frequency counts across the *answering attention questions* (η^2 = .212) category indicated a large effect size. The Tests of Within-Subjects effects for the remaining positive behavior categories were nonsignificant.

Table 14

Tests of Within-Subject Effects: Significant Main Effects Across Positive Behavior Categories

Effect		SS	df	MS	F	*p*	Partial η^2
Answering Attention Questions							
	Time	7078.424	5	1415.685	2.694	.031*	.212
	Error (Time)	26270.765	50	525.415			
Following Activity Rules							
	Time	6389.159	3.504[1]	1823.402	5.612	<.001**	.132
	Error (Time)	42123.830	129.647[1]	324.911			
Contributing to Group Discussion							
	Time	8.976	5.000[2]	1.795	2.947	.014*	.074
	Error (Time)	112.683	143.520[1]	.785			
Helping a Peer							
	Time	8.472	3.356[1]	2.525	2.860	.034*	.072
	Error (Time)	109.596	124.159[1]	.883			

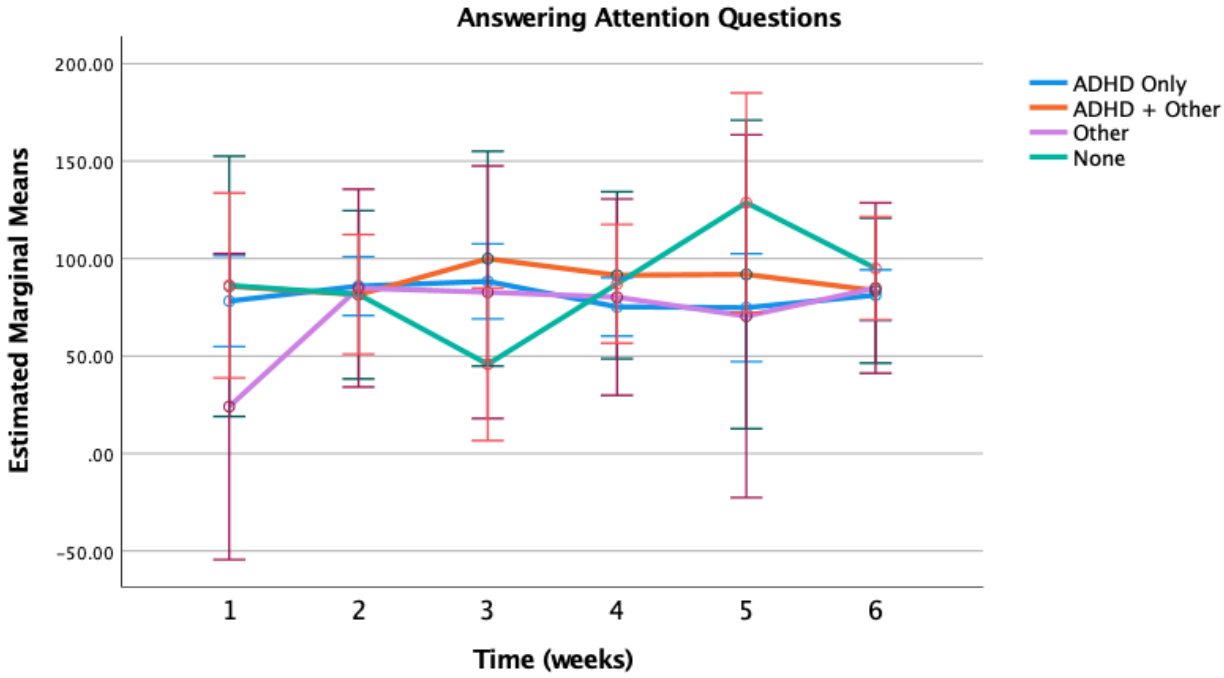
Note. [1]Greenhouse-Geisser correction applied; [2]Huynh-Feldt correction applied. SS = sum of squares; df = degrees of freedom, MS = mean square. Partial eta squared (η^2) values of 0.01 indicates a small effect; values of 0.06 indicate a medium effect; and values of 0.14 indicate a large effect.

Polynomial contrasts revealed significant linear trends for *answering attention questions* (F (1, 10) = 5.54; p = .040; η^2 = .356) and *contributing to a group discussion* (F (1, 37) = 5.44; p = .025; η^2 = .128) positive behavior categories. The linear trend within the *answering attention questions* category indicated a large effect size, and the linear trend within the *contributing to a group discussion* category indicated a medium effect size. The linear trend for *helping a peer* approached statistical significance (F (1, 37) = 3.95; p = .059; η^2 = .097). A quadratic trend, with a large effect size, was indicated for the *following activity rules* (F (1, 37) = 18.13; p < .001; η^2 = .329) behavior category. The graphic representations of the estimated

marginal means of each of these four positive behavior categories over time are presented below (Figures 8–13).

Figure 8

Estimated Marginal Means Across Time: Answering Attention Questions by Group

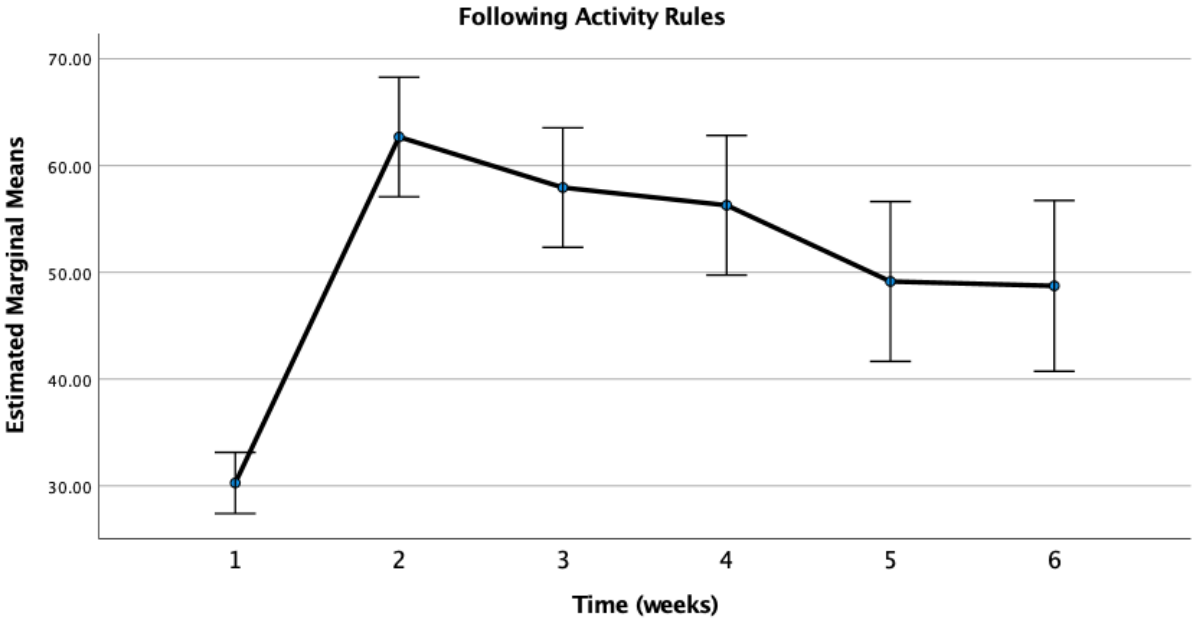


Covariates appearing in the model are evaluated at the following values: BPM_score = 16.8889

Error bars: 95% CI

Figure 9

Estimated Marginal Means Across Time: Following Activity Rules



Covariates appearing in the model are evaluated at the following values: BPM_score = 17.2292

Error bars: 95% CI

Figure 10

Estimated Marginal Means Across Time: Following Activity Rules by Group

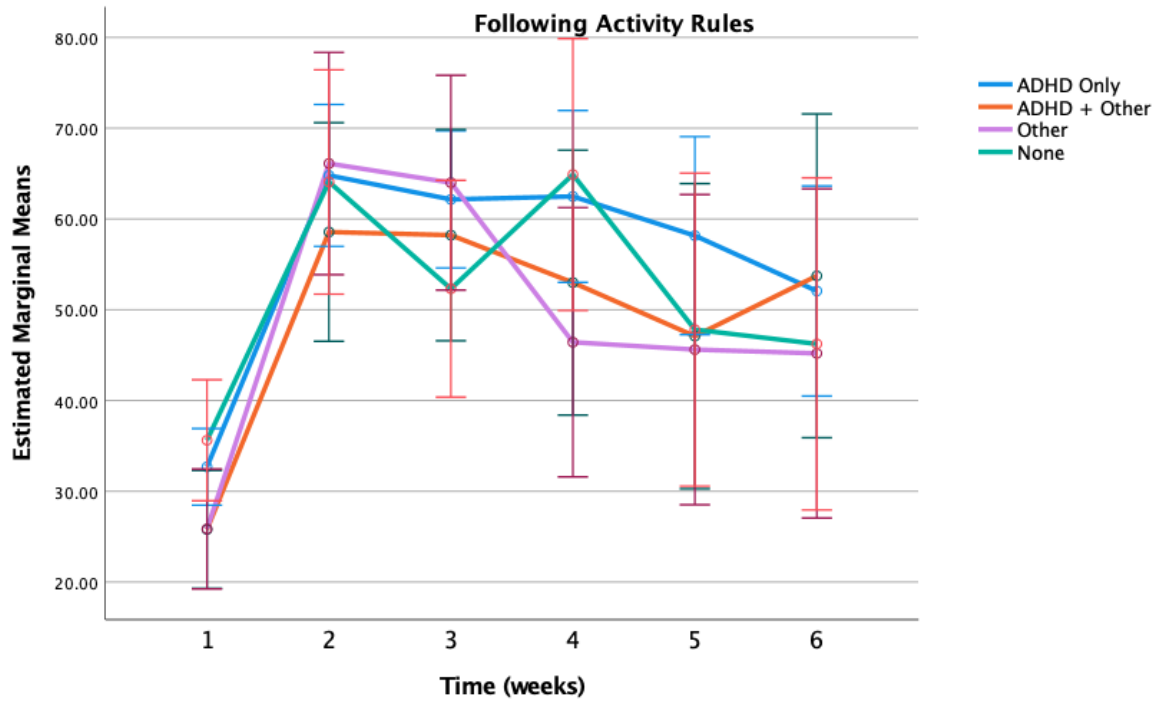


Figure 11

Estimated Marginal Means Across Time: Contributing to Group Discussion

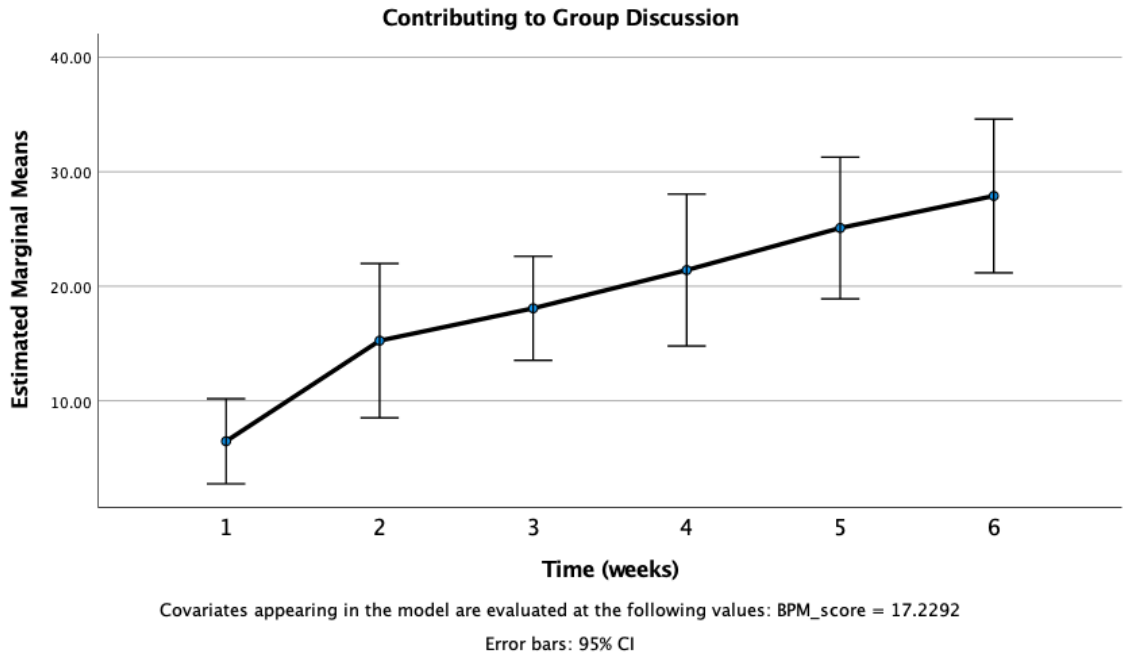
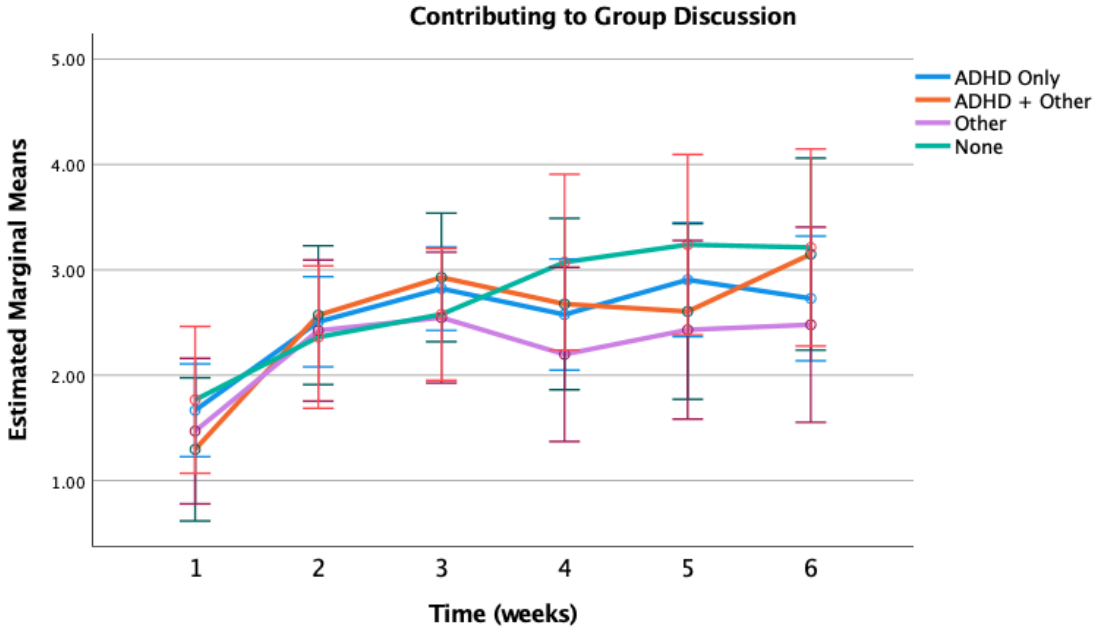


Figure 12

Estimated Marginal Means Across Time: Contributing to Group Discussion by Group

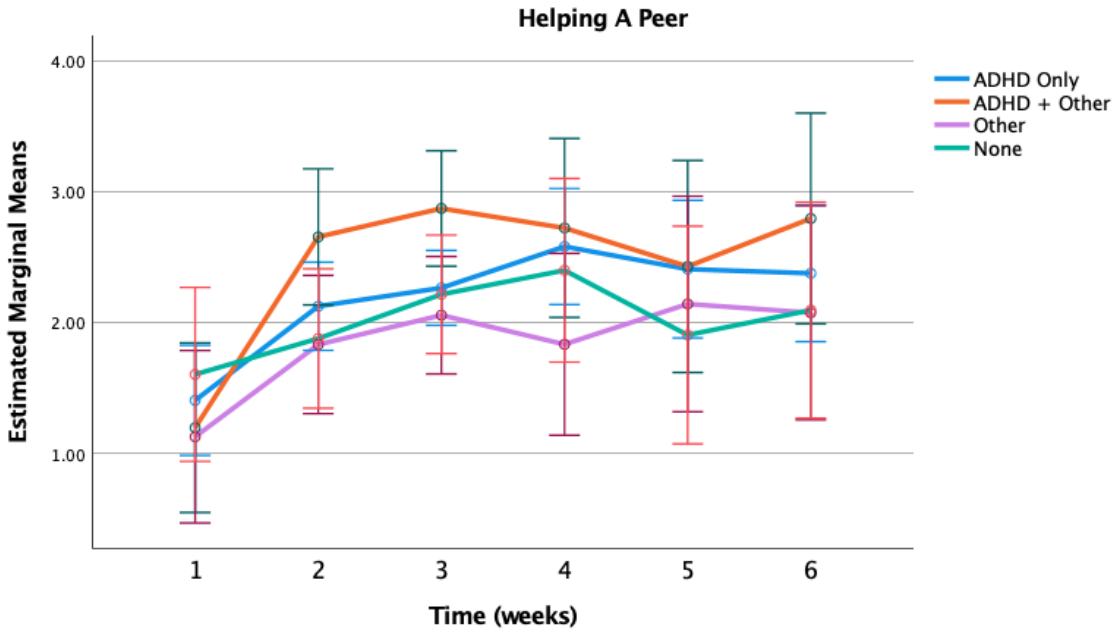


Covariates appearing in the model are evaluated at the following values: BPM_score = 17.2667

Error bars: 95% CI

Figure 13

Estimated Marginal Means Across Time: Helping a Peer



Covariates appearing in the model are evaluated at the following values: BPM_score = 17.2667

Error bars: 95% CI

Adjusted means indicated how mean frequency counts of *answering attention questions*, *following activity rules*, *contributing to group discussion,* and *helping a peer* behaviors varied over the six week program (Table 15). Post-hoc tests were then conducted to identify the week(s) in which there were significant differences in mean frequency scores for the *answering attention questions*, *following activity rules, contributing to group discussion,* and *helping a peer* behavior categories (Table 16).

Table 15

Estimated Marginal Means: Main Effect of Time, Positive Behavior Categories

Behavior Category	Time (week)	*M*	*SE*	95% CI	
				Lower Bound	Upper Bound
Answering Attention Questions[1]	1	68.63	12.99	39.68	97.58
	2	83.52	8.40	64.81	102.23
	3	79.25	10.73	55.36	103.15
	4	83.54	8.34	64.96	102.12
	5	91.51	15.41	57.18	125.83
	6	82.21	7.23	70.09	102.33
Following Activity Rules[2]	1	30.00	1.50	26.96	33.04
	2	63.40	2.78	57.76	69.03
	3	59.18	2.69	53.73	64.62
	4	56.71	3.37	49.87	63.54
	5	49.68	3.88	41.81	57.55
	6	49.32	4.12	40.97	57.67
Contributing to Group Discussion[2]	1	1.55	0.16	1.23	1.87
	2	2.47	.015	2.16	2.78
	3	2.72	0.14	2.43	3.01
	4	2.63	0.18	2.25	3.01
	5	2.80	0.19	2.41	3.19
	6	2.89	0.21	2.47	3.32

Table 15 (continued)

Helping a Peer[2]	1	1.33	0.15	1.03	1.64
	2	2.12	0.12	1.88	2.37
	3	2.35	0.10	2.15	2.56
	4	4.39	0.16	2.06	2.71
	5	2.22	0.19	1.84	2.60
	6	2.34	0.19	1.96	2.71

Note. Covariates appearing in this model were evaluated at a BPM value of [1]16.8889 or [2]17.2667.

Table 16

Post-Hoc Pairwise Comparisons: Time

	Mean Difference	*SE*	Bonferroni-Holm Adjusted *p* value
Following Activity Rules			
Week 1 to Week 2	33.40	2.83	< .001**
to Week 3	29.18	2.86	< .001**
to Week 4	26.71	3.02	< .001**
to Week 5	19.68	4.22	.001*
to Week 6	19.32	4.48	.003*
Week 2 to Week 5	-13.72	3.78	.017*
to Week 6	-14.08	4.10	.027*
Contributing to Group Discussion			
Week 1 to Week 2	0.92	0.13	< .001**
to Week 3	1.17	0.18	< .001**
to Week 4	1.08	0.16	< .001**
to Week 5	1.25	0.23	< .001**
to Week 6	1.34	0.23	< .001**
Helping a Peer			
Week 1 to Week 2	0.79	0.13	< .001**
to Week 3	1.02	0.15	< .001**
to Week 4	1.05	0.19	< .001**
to Week 5	0.89	0.24	.014*
to Week 6	1.00	0.24	.005*

Note. *p < 0.05; **p < .001. *SE* = standard error.

Bonferroni-Holm post-hoc tests revealed nonsignificant differences in mean frequency counts over time for the *answering attention questions* behavior category. The *following activity rules*, *contributing to group discussion*, and *helping a peer* mean frequency counts were significantly different between week 1 and all consecutive time points. Post-hoc pairwise comparisons also indicated a significant difference in mean frequency counts of *following activity rules* between week 2 and week 5 ($p = .017$) and week 2 and week 6 ($p = .027$).

Post-hoc multivariate analyses were also conducted for those behavior categories that demonstrated significant main effects of time but were previously identified as having violated the sphericity assumption (Table 17). The multivariate repeated measures ANCOVA for the within-subjects effect of time indicated significant overall differences in, and large effect sizes for, the mean frequency counts of *following activity rules* ($F(5, 33) = 28.07$; $p < .001$; $\eta^2 = .810$), *contributing to group discussions* ($F(5, 33) = 13.21$; $p < .001$; $\eta^2 = .667$), and *helping a peer* ($F(5, 33) = 9.42$; $p < .001$; $\eta^2 = .588$) behavior categories across the weeks of treatment.

Table 17

Post-Hoc Multivariate ANCOVA: Positive Behavior Categories

Within-Subjects Effect	Value	F	Hypothesis df	Error df	*p*	Partial η^2
Time						
Following Activity Rules	.810	28.066[1]	5.00	33.00	<.001**	.810
Contributing to Group Discussion	.667	13.206[1]	5.00	33.00	<.001**	.667
Helping a Peer	.588	9.420[1]	5.00	33.00	<.001**	.588

Note. [1]Pillai's trace.

The results of the Tests of Between-Subjects Effects indicated significant main effects for group across the *complying with counselor commands* ($p = .005$; $\eta^2 = .290$), *ignoring a negative stimulus* ($p = .039$; $\eta^2 = .200$), *helping a peer* ($p = .035$; $\eta^2 = .205$), and *sharing with a peer* ($p < .001$; $\eta^2 = .442$) behavior categories (Table 18). Results also indicated significant interactions between group and the covariate or *complying with a command* ($p < .001$; $\eta^2 = .362$) and *sharing with a peer* ($p < .001$; $\eta^2 = .383$), behavior categories.

Table 18

Tests of Between-Subject Effects: Significant Main Effects and Interactions of Positive Behavior Categories

Source	SS	df	MS	F	*p*	Partial η^2
Covariate						
Following Activity Rules	7121.927	1	7121.927	6.694	.014*	.153
Group						
Complying with a Command	33.126	3	11.042	5.035	.005*	.290
Ignoring a Negative Stimulus	3086.934	3	1028.978	3.083	.039*	.200
Helping a Peer	20.969	3	6.990	3.180	.035*	.205
Sharing with a Peer	132.028	3	44.009	9.769	<.001**	.442
Covariate*Group						
Complying with a Command	45.945	3	15.315	6.983	<.001**	.362
Sharing with a Peer	103.354	3	34.451	7.647	<.001**	.383

However, post-hoc univariate repeated measures ANOVA tests were nonsignificant for *complying with a command* ($F(3, 37) = 0.97$; $p = .416$), *ignoring a negative stimulus* ($F(3, 37) = 1.11$; $p = .356$), *helping a peer* ($F(3, 37) = 1.44$; $p = .246$), and *sharing with a peer* ($F(3, 37) = 2.70$; $p = .060$) behavior categories, countering the false positive demonstrated by the significant F tests seen in Table 18.

To further examine the relationship between the covariate and the *following activity rules* behavior, a secondary analysis was performed. A linear regression analysis, with covariate as the independent variable and the *following activity rules* behavior across the weeks of treatment as the dependent variable, indicated three statistically significant predictive relationships of the covariate on the frequency of *following activity rules*. For every unit (point) increase in the total BPM score at pre-treatment, the following changes in frequency of *following activity rules* was predicted: a 1.05 decrease at week 2 (F (1, 46) = 5.926, p = .019, R^2 = .095), a 1.14 decrease at week 3 (F (1, 46) = 6.858, p = .012, R^2 = .111), and a 1.13 decrease at week 4 (F (1, 46) = 4.232, p = .012, R^2 = .064). The covariate yielded nonsignificant effects as a predictor of *following activity rules* across all remaining weeks.

Table 19

Linear Regression Analysis: Covariate Predicting Following Activity Rules Behavior

Predictor (*DV*)	B	95% CI	β	*t*	*p*	Adjusted R^2
BPM (*FAR Week 1*)	-.451	[-.971, .070]	-.249	-1.743	.088	.042
BPM (*FAR Week 2*)	-1.048	[-1.914, -.181]	-.338	-2.434	.019*	.095
BPM (*FAR Week 3*)	-1.140	[-2.016, -.264]	-.360	-2.619	.012*	.111
BPM (*FAR Week 4*)	-1.131	[-2.238, -.24]	-.290	-2.057	.045*	.064
BPM (*FAR Week 5*)	-.972	[-2.197, .253]	-.229	-1.597	.117	.032
BPM (*FAR Week 6*)	-.768	[-2.027, .492]	-.178	-1.226	.226	.011

Note. BPM = pre-treatment BPM total score. FAR = following activity rules. CI = confidence interval for B.

Summary of Results. The repeated measures ANCOVA tests indicated a statistically significant difference in mean frequency counts of behavior over time across three positive

behavior categories: *following activity rules*, *contributing to group discussion*, and *helping a peer*. These differences were characterized by medium effect sizes. Polynomial contrasts indicated a significant quadratic trend, of a large effect size, in frequency of *following activity rules* over time. A significant linear trend, of medium effect size, was observed in frequency of *contributing to group discussion* behaviors over time. Post-hoc analyses indicated the greatest difference in mean frequency count of *following activity rules* occurred between week 1 and week 2 (+33.40; $p < .001$) and that there was a significant decrease in mean frequency count for *following activity rules* between week 2 and week 6 (-14.08; $p = .027$).

The repeated measures ANCOVA analyses indicated mean frequency counts for following activity rules were statistically significantly different based on parent-reported score of child functioning at baseline (covariate; BPM total score). These analyses also indicated statistically significant differences in mean frequency counts across concurrent pharmacotherapy groups for four positive behavior categories: *complying with a command*, *ignoring a negative stimulus*, *helping a peer*, and *sharing with a peer*. However, post-hoc univariate tests were nonsignificant for group effects across these behavior categories. This suggests that the significant between-subjects omnibus tests were false positives.

Research Question 3

Do children who complete the 6-week STP show decreases in frequency of negative behaviors, including: interrupting, whining, swearing, teasing, verbal abuse, leaving the activity area, intentional aggression, and intentional destruction?

Hypothesis 3. The frequency of behaviors within each negative behavior category will decrease throughout the 6-week STP.

Results. To test this hypothesis, a repeated measures ANCOVA model with concurrent pharmacotherapy group as the between-subjects factor, time as the within-subjects factor, and the pre-treatment BPM total score as the covariate was used. The analysis included 48 participants. Separate repeated measures ANCOVA models were used to examine change in frequency counts across each of the eight negative behavior categories on the STP point system. Case analyses revealed outlying values, within the *swearing*, *verbal abuse*, *leaving the activity area*, *intentional aggression*, and *intentional destruction* negative behavior categories (Table 20).

Table 20

Case Analyses: Standardized Residuals Across Negative Behavior Categories

ID	BPM Score	Group	Standardized Residual					
			Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Swearing								
15	17	Other	4.85					4.48
27	26	None	3.55					
39	27	ADHD Only			6.58	5.45	6.28	
50	--	None		3.97				
Verbal Abuse								
19	8	ADHD + Other		4.99				
21	17	ADHD + Other		3.71	4.37	4.06	5.51	3.73
27	26	None				4.37		
36	17	None				4.06		3.97
39	27	ADHD Only	4.73					
Leaving the Activity Area								
36	17	None	4.34	6.46	6.59	6.17	6.32	5.48

Table 20 (continued)

Intentional Aggression								
7	16	ADHD Only		3.57				
8	17	ADHD Only			5.04			
27	26	None		3.57				
36	17	None	5.34			5.96	6.78	6.95
Intentional Destruction								
6	13	Other	4.00		4.69			
8	17	ADHD Only	4.00					
22	9	ADHD + Other	4.00					
36	17	None		5.89		6.39	6.78	6.09

A sensitivity analysis was completed to determine the impact of outlying values on the ANCOVA tests results (Table 21). In the absence of the outlier in the *intentional destruction* group, the ANCOVA yielded a significant main effect ($p = .004$), whereas both main effects were nonsignificant when the outlier was included. This discrepancy indicated that the outlying cases in the *intentional destruction* behavior category demonstrated excessive influence on the results. Thus, those outliers were omitted from the data set used to run the primary ANCOVA analysis.

Table 21

Sensitivity Analyses: Main Effects and Interaction

		Tests of Between-Subjects Effects		Excessive Influence?
		Included	Excluded	
Source		*p*	*p*	
Swearing				No
	Group	.451	.390	
	BPM	.080	.192	
	Group*BPM	.438	.266	

Table 21 (continued)

Verbal Abuse			**No**
	Group	.182	.893	
	BPM	.273	.545	
	Group*BPM	.304	.842	
Leaving the Activity Area			No
	Group	.907.	.261	
	BPM	.664	.644	
	Group*BPM	.841	.388	
Intentional Aggression			No
	Group	.894	.366	
	BPM	.407	.886	
	Group*BPM	.539	.349	
Intentional Destruction			Yes
	Group	.944	.176	
	BPM	.647	.004*	
	Group*BPM	.782	.094	

Note. *p < 0.05; **p < .001

Levene's test was used to determine homogeneity of variances for mean frequency counts within each of the eight negative behavior categories in the STP point system. The results indicate that the error variance of mean frequency counts was significantly different across the weeks of treatment in each of the negative behavior categories, with the exception of *interrupting* (*p* values nonsignificant across time) (Table 22). To address the violation in the homogeneity of variance assumption across these seven negative behavior categories, a log transformation was applied. In cases where a dependent variable had a value of '0', a '1' was added to the score before the log transformation was applied.

Table 22

Levene's Test: Negative Behavior Categories

Variable		F	df1	df2	Sig.
Whining					
	Week 2	2.926	3	44	.044*
	Week 3	8.816	3	44	<.001**
	Week 5	5.296	3	44	.003*
	Week 6	3.568	3	44	.021*
Swearing					
	Week 1	11.888	3	44	<.001**
	Week 2	12.323	3	44	<.001**
Teasing					
	Week 1	3.420	3	44	.025*
	Week 2	5.335	3	44	.003*
	Week 5	3.180	3	44	.033*
	Week 6	2.878	3	44	.047*
Verbal Abuse					
	Week 2	17.285	3	44	<.001**
	Week 3	7.044	3	44	<.001**
	Week 4	5.277	3	44	.003*
	Week 5	4.895	3	44	.005*
	Week 6	3.161	3	44	.034*
Leaving the Activity Area					
	Week 2	5.726	3	44	.002*
	Week 3	6.073	3	44	.001*
	Week 4	5.952	3	44	.002
	Week 5	5.142	3	44	.004*
Intentional Aggression					
	Week 4	4.663	3	44	.006*
	Week 5	6.602	3	44	<.001**
	Week 6	7.762	3	44	<.001**

Table 22 (continued)

Intentional Destruction					
	Week 2	6.171	3	44	.001*
	Week 3	4.512	3	44	.008*
	Week 4	6.364	3	44	.001*
	Week 5	6.706	3	44	<.001**
	Week 6	5.805	3	44	.002*

The primary repeated measures ANCOVA was then conducted. Mauchly's *W* was significant for seven behavior categories, indicating that the assumption of sphericity had been violated (Table 23). Based on the estimated sphericity (ε), the degrees of freedom was corrected using either the Greenhouse-Geisser or Huynh-Feldt correction.

Table 23

Mauchly's Test: Negative Behavior Categories

Behavior Category	Mauchly's *W*	*p*	Epsilon (ε)	
			Greenhouse-Geisser	Huynh-Feldt
Interrupting[1]	.090	<.001	.560	.712
Whining[1]	.414	.002	.691	.896
Swearing[1]	.382	<.001	.738	.964
Verbal Abuse[1]	.503	.025	.788	1.000
Leaving the Activity Area[1]	.203	<.001	.601	.769
Intentional Aggression[1]	.372	,.001	.725	.946
Intentional Destruction[1]	.142	<.001	.560	.712

Note. [1]Greenhouse-Geisser correction indicated. [2]Huynh-Feldt correction indicated.

The primary repeated measures ANCOVA was then conducted. The results of the Tests of Within-Subjects Effects yielded nonsignificant results across all negative behavior categories, indicating mean frequency counts of the *interrupting*, *whining*, *swearing*, *teasing*, *verbal abuse*,

leaving the activity area, *intentional aggression*, and *intentional destruction* behavior categories did not significantly differ across the course of the six week program. The results of the Tests of Between-Subjects Effects indicated a significant main effect for the covariate in relation to the *swearing* ($p = .046$; $\eta^2 = .096$) behavior category (Table 24). The Tests of Between-Subjects Effects for the remaining seven behavior categories yielded nonsignificant effects.

Table 24

Tests of Between-Subject Effects: Main Effects and Interactions of Negative Behavior Categories

Source		SS	df	MS	F	*p*	Partial η^2
Covariate							
	Swearing	6.136	1	6.136	4.253	.046*	.096

To further examine the relationship between the covariate and *swearing* behaviors, a secondary analysis was performed. A linear regression analysis, with covariate as the independent variable and *swearing* behavior across the weeks of treatment as the dependent variable, indicated one statistically significant predictive relationship of covariate on *swearing* behavior outcomes (F (1, 46) = 4.203, $p = .046$, $R^2 = .084$) (Table 25): for every unit (point) increase in the total BPM score at pre-treatment, a .041 increase in *swearing* frequency was predicted at week 5 ($t = 2.05$, $p = .046$). The covariate yielded nonsignificant effects as a predictor of *swearing* behavior across all other weeks.

Table 25

Linear Regression Analysis: Covariate Predicting Swearing Behavior

Predictor (*DV*)	B	95% CI	β	*t*	*p*	Adjusted R^2
BPM (*Swearing Week 1*)	.015	[-.004, .034]	.224	1.562	.125	.030

Table 25 (continued)

BPM (*Swearing Week 2*)	.022	[-.015, .059]	.172	1.184	.242	.008
BPM (*Swearing Week 3*)	.024	[-.006, .054]	.234	1.629	.110	.034
BPM (*Swearing Week 4*)	.025	[-.021, .071]	.157	1.078	.287	.003
BPM (*Swearing Week 5*)	.041	[.001, .082]	.289	2.050	.046*	.064
BPM (*Swearing Week 6*)	.031	[-.004, .066]	.257	1.800	.078	.046

Note. BPM = pre-treatment BPM total score. CI = confidence interval for B.

Summary of Results. The repeated measures ANCOVA tests indicated nonsignificant differences in mean frequency counts of all seven negative behavior categories over time. As such, there was insufficient evidence to support the hypothesis that the frequency of behaviors within each negative behavior category decrease throughout the 6-week STP. The Tests of Between-Subject Effects revealed mean frequency counts for *swearing* were statistically significantly different based on parent-reported scores of child functioning at baseline (covariate; BPM total score). Although this effect was not of primary interest, it is notable because it indicated the covariate is a statistically significant predictor for the *swearing* behavior category. Results of the linear regression revealed that for every one point increase in the BPM total score at pre-treatment, the mean frequency count of *swearing* behavior at week 5 is predicted to increase by .041 (Table 25). The BPM total score was a nonsignificant predictor of *swearing* behavior mean frequency counts at all other time points.

Research Question 4

Do children who complete the 6-week STP show improvements in academic learning center outcomes, as demonstrated by total points earned for: following learning center rules

(Behavior score), completing homework assignments and for completing assignments with 80% or higher accuracy (Academic score), and other positive behaviors (bonus points).

Hypothesis 4. The points earned for positive behaviors (Behavior score plus Bonus points) and homework completion and accuracy (Academic score) will increase throughout the 6-week STP.

Results. To test this hypothesis, a repeated measures ANCOVA model with mean ALC total scores as the dependent variable, concurrent pharmacotherapy group as the between-subjects factor, time as the within-subjects factor, and pre-treatment BPM total score as the covariate was used. The analysis included 48 participants. Case analysis did not reveal any standardized residuals with absolute values larger than 3. Levene's test was used to examine homogeneity of variances for mean ALC total scores across the six weeks of treatment. Results indicated the error variance of the mean ALC total scores was similar across groups. A preliminary Test of Between-Subject Effects indicated the interaction term was not significant (p = .454), meaning the regression slopes for the covariate did not differ between groups, and the homogeneity of regression slopes assumption was met.

The repeated measures ANCOVA model was then run without the interaction term. Mauchly's W test was significant (p = .005), indicating that the sphericity assumption had been violated. The Greenhouse-Geisser estimated ε values was greater than 0.75 (ε = 0.79); therefore the Huynh-Feldt degrees of freedom correction was applied to the subsequent Tests of Within-Subjects Effects (Table 26). The results indicated a significant difference in mean ALC

total scores over time ($p < .001$; $\eta^2 = .125$) with a medium effect size. Adjusted means indicated how mean ALC total scores varied across the six week program (Table 27).

Table 26

Tests of Within-Subject Effects: ALC Total Score

Effect	SS	df[1]	MS	F	*p*	Partial η^2
ALC Total						
Time	1,279,345.	5.00	255,869.11	5.6	<.001**	.125
Error (Time)	8,992,068.	200.0	44,960.345			

Note. [1]Huynh-Feldt correction applied when Mauchly's W test yielded a significant p-value.

Table 27

Estimated Marginal Means: Main Effect of Time on Mean ALC Total Scores

Time (week)	*M*	*SE*	95% CI	
			Lower Bound	Upper Bound
1	269.94	16.91	235.75	304.12
2	696.48	44.22	606.11	784.86
3	787.93	35.64	715.90	859.96
4	816.82	43.19	729.53	904.11
5	641.35	49.98	540.34	742.37
6	660.50	54.71	549.93	771.08

Note. Covariates appearing in this model were evaluated at a BPM value of 17.2292.

Bonferroni-Holm post-hoc tests revealed significant increases in mean ALC total scores between week 1 and all consecutive time points (Table 28). Results also indicated significant decreases in ALC total scores between week 4 and week 5 ($p = .022$).

Table 28

Post-Hoc Pairwise Comparisons: Significant Differences in ALC Total Scores Across Time

	Mean Difference	Std. Error	Bonferroni-Holm Adjusted *p* value
Week 1 to Week 2	425.54	41.99	<.001**
to Week 3	517.99	35.37	<.001**
to Week 4	546.88	36.10	<.001**
to Week 5	371.41	51.5	<.001**
to Week 6	390.56	56.45	<.001**
Week 4 to Week 5	-175.47	49.90	.022*

Note. Positive mean difference values indicate increase in mean score over given weeks; negative mean difference value indicates decrease in mean score over given week. * $p < 0.05$; ** $p < .001$.

A post-hoc multivariate analysis was also conducted because the dependent variable, while demonstrating a significant main effect of time, was previously identified as having violated the sphericity assumption. The multivariate repeated measures ANCOVA for the within-subjects effect of time indicated a significant overall difference in, and a large effect size for, the mean ALC total score ($F(5, 36) = 60.72$; $p < .001$; $\eta^2 = .894$) across the weeks of treatment (Table 29).

Table 29

Post Hoc Multivariate ANCOVA: ALC Total Score

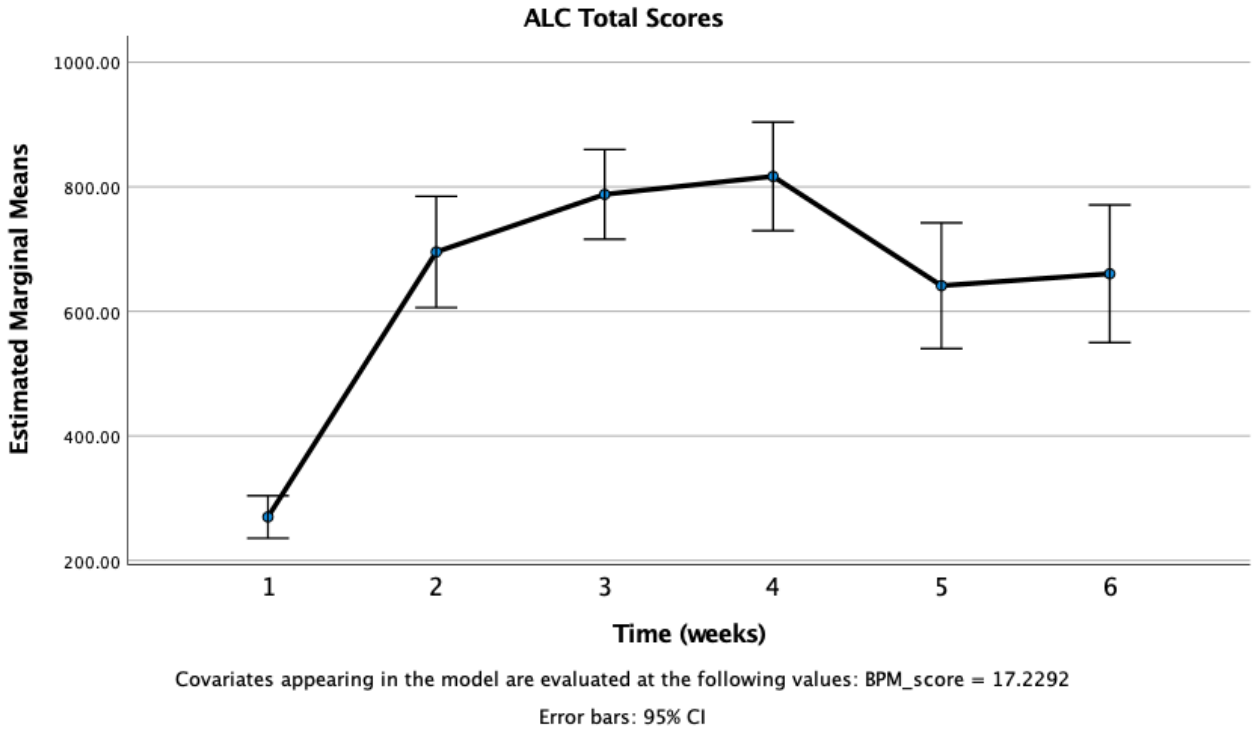
Within-Subjects Effect	Value	F	Hypothesis df	Error df	*p*	Partial η^2
Time	.894	60.718[1]	5.00	36.00	<.001**	.894

Note. [1]Pillai's Trace.

Polynomial contrasts revealed a significant quadratic trend, with a large effect size, for ALC total scores over time ($F(1, 40) = 15.89$; $p < .001$; $\eta^2 = .284$). The graphic representation of the estimated marginal means of ALC scores over time are presented below (Figure 13).

Figure 14

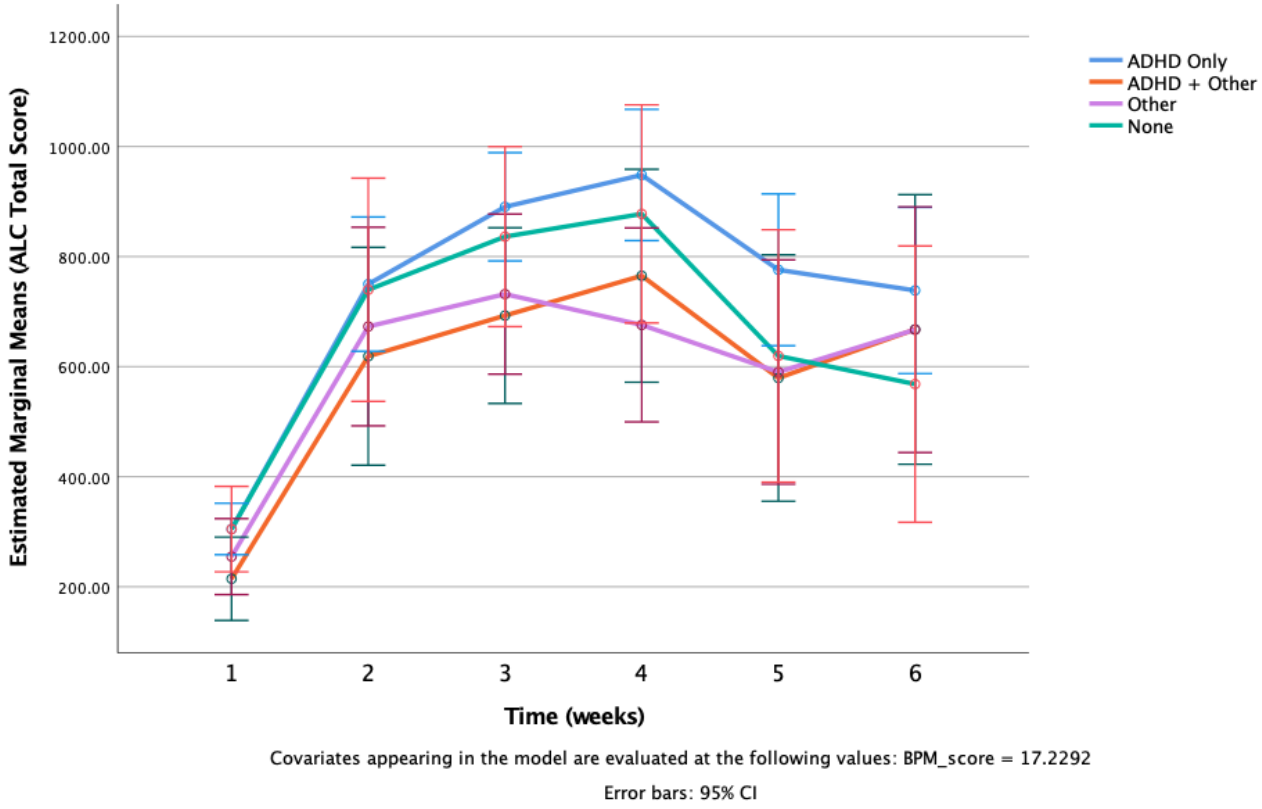
Estimated Marginal Means Across Time: ALC Total Scores



The results of the Test of Between-Subjects Effects indicated a nonsignificant main effect for group ($p = .488$) suggesting similar mean ALC total scores across concurrent pharmacotherapy groups over time (Figure 15).

Figure 15

Mean ALC Total Scores as a Function of Week and Pharmacotherapy Group



Summary of Results. There was evidence of statistically significant differences in mean ALC total scores over time, however these scores did not increase across each week as predicted. Post-hoc analyses indicated one instance of a significant decrease in mean ALC total scores between week 4 and week 5 (-175.47; $p = .022$). Polynomial contrasts revealed a significant quadratic trend, corresponding with a large effect size, in ALC total scores over time. The repeated measures ANCOVA also yielded nonsignificant differences in mean ALC total scores between concurrent pharmacotherapy groups, suggesting participants across each

group had similar patterns of change in mean ALC total scores over the course of the 6-week STP .

Research Question 5

Do children who complete the 6-week STP show improvements in the percentage of DRC target behaviors achieved?

Hypothesis 5. The percentage of DRC target behaviors achieved will increase throughout the 6-week STP.

Results. To test the final hypothesis, a repeated measures ANCOVA model with mean DRC scores as the dependent variable, concurrent pharmacotherapy group as the between-subjects factor, time as the within-subjects factor, and pre-treatment BPM total score as the covariate was used. The analysis included 37 participants. Case analysis did not reveal any standardized residuals with absolute values larger than 3. Levene's test did not yield any significant differences between error variance of the mean DRC scores across the six weeks of treatment. Thus, the homogeneity of variance assumption was met. A preliminary Test of Between-Subject Effects indicated the interaction term was not significant meaning the regression slopes for the covariate did not differ between groups, and the homogeneity of regression slopes assumption was met.

The repeated measures ANCOVA model was then run without the interaction term. Mauchly's W test was not significant ($p = .159$) indicating the sphericity assumption was met. The results of the Levene's test indicated there were no significant differences in the variances of the mean DRC total scores across groups. The Test of Within-Subjects Effects yielded

nonsignificant main effects of time (p = .645) and nonsignificant interaction effects between time and concurrent pharmacotherapy group (p = .528). The Tests of Between-Subjects Effects also yielded nonsignificant main effects of group (p = .574). On average, children experienced the same degree of improvement (6.13%) in daily goal attainment over the course of the 6-week STP (75.16 – 81.29%) (Figure 16)

Figure 16

Estimated Marginal Means Across Time: DRC Total Scores

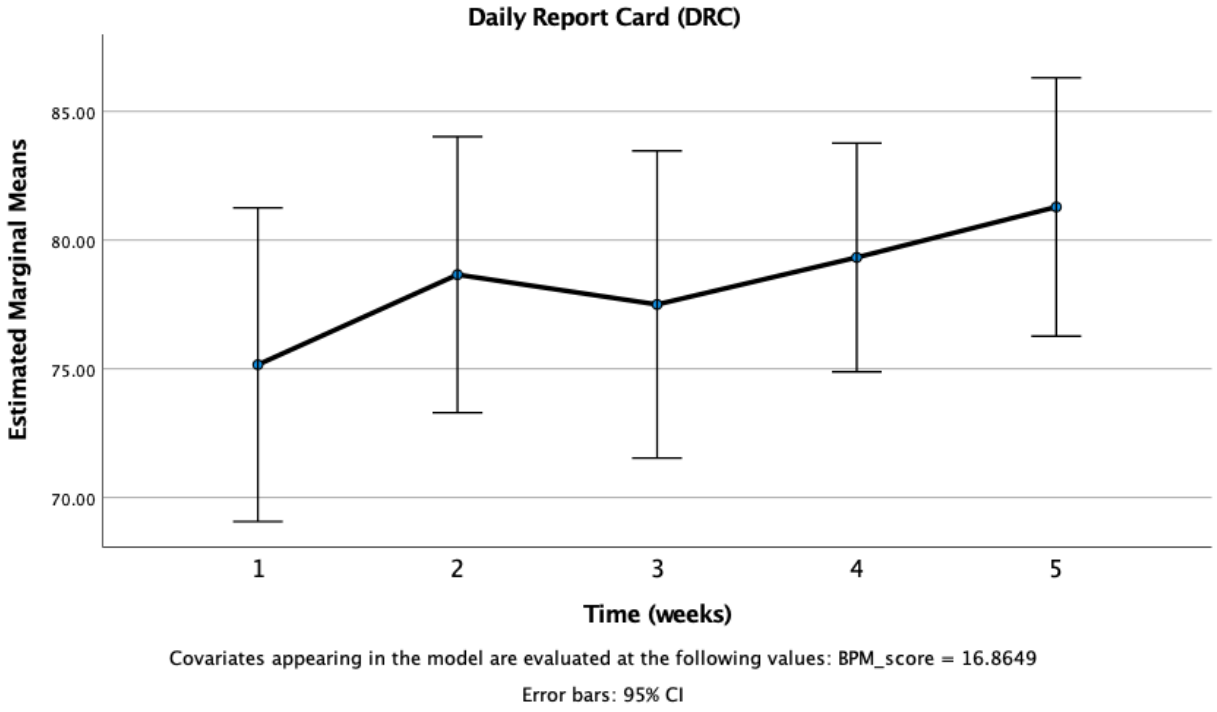
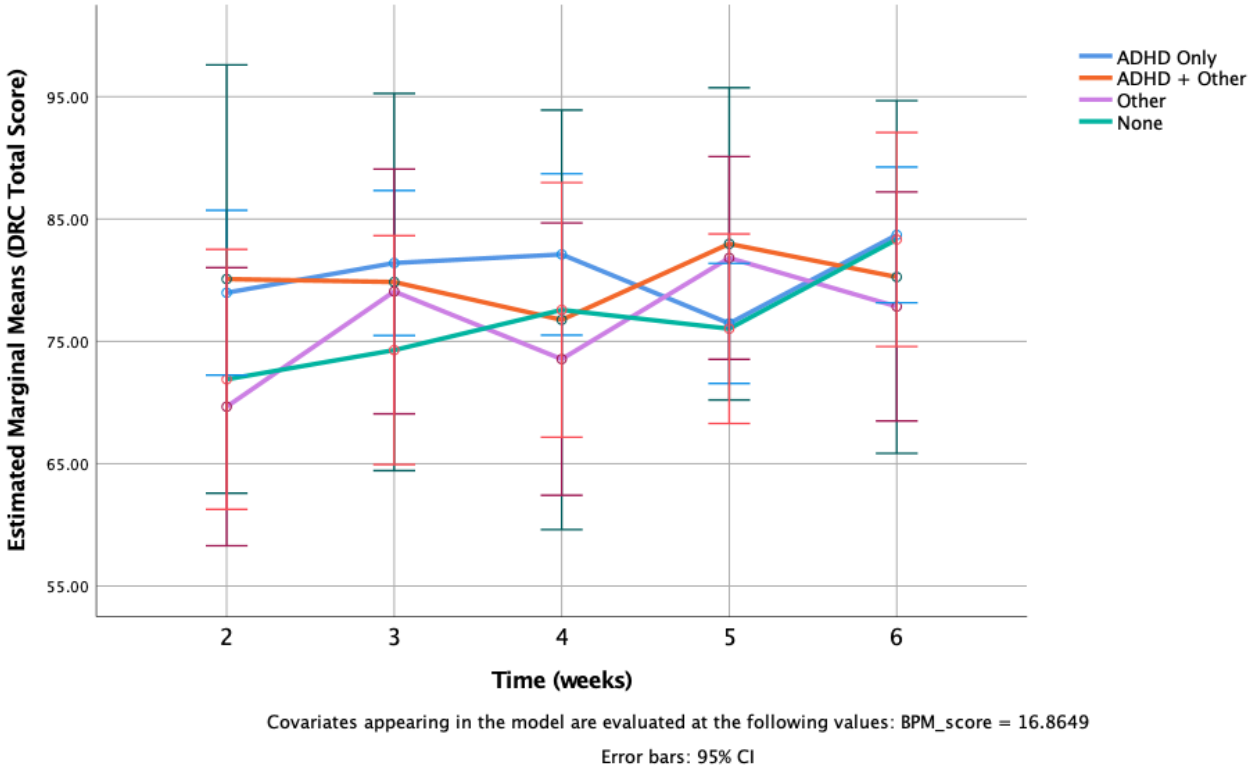


Figure 17

Mean DRC Total Scores as a Function of Week and Pharmacotherapy Group



Summary of Results. The repeated measures ANCOVA analysis indicated nonsignificant differences in mean DRC total scores, both across time and between groups. Although there were time points where mean DRC total scores fluctuated, these differences in scores were not statistically significantly different. Moreover, each concurrent pharmacotherapy group demonstrated similar patterns of change in mean DRC total scores over time (Figure 17). In all, there was insufficient evidence to support the hypothesis that mean DRC total scores would increase over the course of the 6-week STP.

Chapter V: Discussion

Overview of Findings

This is the first study to investigate the effectiveness of a 6-week adaptation of the STP delivered in a community setting, a study made even more unique by defining behavioral improvement in terms of socially valid target behaviors that are meaningful to both providers and families. Study findings contribute to the existing treatment outcome literature by extending upon what is known about the clinical utility of the STP model. Study findings also raise important considerations about designing and implementing more rigorous studies of the various STP models.

Program Effectiveness

A program effectiveness study requires clearly defined program objectives, each with a corresponding outcome measure (Deniston et al., 1968). The first STP was established with the objectives of improving children's behavior and social functioning and children's sense of self-efficacy (Pelham & Hoza, 1996). The program objectives for the 6-week STP were similar: improving children's behavioral and social functioning, as measured by behavioral rating scales (parent-report) and behavioral observations (counselor-report). The program effectiveness of the 6-week STP was evaluated using both group-based, average outcomes and, when possible, rates of clinically reliable change (improvement; deterioration) subsequent to the intervention.

The first aim of this study was to evaluate the effects of the adapted STP in relation to parent-reported child functioning across the domains of: inattention, hyperactivity/impulsivity, learning problems, executive functioning, defiance/aggression, and peer relations. Although

there were significant improvements in children's functioning across the inattention, hyperactivity/impulsivity, and peer relations domains, these improvements were not clinically significant because the average scores remained in the *clinical* range ($T \geq 65$; [69.60 – 71.28]). At post-treatment, there was also evidence of significant differences in parent-reported scores of child functioning in the domain of learning problems based on children's concurrent pharmacotherapy status. However, these findings are of marginal clinical significance because differences at post-treatment are not significantly different from the pattern of scores obtained at pre-treatment; and because average post-treatment learning problems scores were either on the lower limit of the *sub-clinical* threshold ($T = 60$-64; ADHD Only, 60.00; Other, 61.44) or remained below the threshold of significance ($T = 40$-59; ADHD + Other, 50.00; None, 50.13).

These findings are the result of statistical analyses commonly implemented in program evaluation studies, and generally lend themselves to assessing whether *treatment as a whole* is effective (Kazdin, 2008). However, there remains the other measure of effectiveness, that is perhaps equally important: *do children experience clinically significant improvement or deterioration in response to treatment?* To this end, an RC analysis was performed. The outcome of interest (parent-reported child functioning at post-treatment) was expressed as one of five change categories: reliable improvement, improvement, no change, deterioration, and reliable deterioration. This, in turn, allowed for *rates* of clinically significant improvement and deterioration to be calculated and compared for each child functioning domain of interest. At post-treatment, there was a significantly higher rate of reliable improvement (45%) than there was rate of reliable deterioration (17.5%) within the hyperactivity/impulsivity domain.

Within the defiance/aggression domain, this statistically significant difference was less stark: 37.5% of children demonstrated improvement with only a slightly lower percentage (32.5%) demonstrating deterioration. In light of the negative correlation between these two parent-reported measures of child functioning at pre-treatment (Table 3), and the high rates of comorbid ODD and ADHD, it is possible that the 6-week STP was effective in improving child functioning that is commonly impaired in children with ADHD (hyperactivity/impulsivity) but insufficient to improve disruptive behaviors (defiance/aggression) that are more characteristic of children with ODD. Thus, children that were otherwise not identified with a comorbid ODD diagnosis may have been found to meet criteria at post-treatment, given a significant difference in presenting problems. Of clinical relevance, the 6-week STP generally yielded greater rates of reliable improvement, compared to rates of deterioration, across each of the six domains of child functioning. In the absence of any prior reports of clinically reliable rates of change subsequent to the STP, these data points are currently of little significance to providers. However, this expression of program effectiveness may be of relevance to families as they calculate the advantages and disadvantages of participating in one treatment over another, or to perhaps forgo treatment altogether.

Program effectiveness was also evaluated when the program objective(s) were defined by: (1) average number of positive and (2) negative behaviors demonstrated, as outlined in the standard STP point system; (3) average total points earned for academics and behavior during ALC; and (4) average percentage of individualized behavioral target goals achieved, as defined by a child's DRC. In this way, these four discrete behavioral outcome measures allowed for

program effectiveness to be calculated using different *units* of functioning, unlike a number of studies to date. Preliminary analyses indicated significant outliers within the following categories of behaviors defined by the STP point system: one outlier for *helping a peer*, three outliers for *sharing with a peer*, four outliers for *swearing*, five outliers for *verbal abuse*, one outlier for *leaving the activity area,* four outliers for *intentional aggression*, and four outliers for intentional destruction. Of note, there were no identifiable patterns of baseline functioning (as defined by the BPM total score) or concurrent pharmacotherapy status among these children with outlying values.

Of the seven positive behavior categories defined by the standard STP point system, children who completed the 6-week STP demonstrated significantly higher frequencies of *following activity rules*, *contributing to group discussion*, and *helping a peer* behaviors over time. Despite the significant linear trend of increasing frequency of *contributing to group discussion* and *helping a peer* behaviors over time, the average increase in behavioral frequency at post-treatment within these two behavior categories may be of relative importance and/or subjective clinical significance to children and their families. For children presenting with symptoms of social anxiety at school, an increase in frequency of *contributing to group discussion* from 1 to 3 over a period of six weeks may prove valuable to academic and social functioning at school. Similarly, for children presenting with high rates of impulsive behaviors which preclude them from attending to the needs of others, an increase in average number of times *helping a peer* demonstrated over a period of six weeks may be a particular marker of success for a child that otherwise receives negative feedback for their inappropriate behaviors.

In other instances, it may be that these prosocial behaviors are of secondary importance to higher base rate behaviors, such as: not completing a task assigned by an adult, engaging in activities without regard for the technical rules, or being easily distracted from the task at hand by extraneous stimuli. To these families, significant increases in *following activity rules* behaviors over the course of the 6-week STP may be of more clinical relevance. The improvement in rate of *following activity rules* behaviors increased linearly from Week 1 to Week 2; however, the overall trend, from Week 1 to Week 6, was significant for a quadratic relationship with a significant downwards trend between Weeks 2 to Weeks 5 and 6. Previous studies have indicated significant improvements in *following activity rules* behaviors subsequent to the 8-week STP (MTA Cooperative Group, 1999; Pelham et al., 2000; Coles et al., 2005; Pelham et al., 2005). However, the observed trends in frequency of this specific behavior category over time is absent from the existing literature. Thus, the quadratic trend in frequency of *following activity rules* behaviors observed in this study are of particular interest. What appears to be a ceiling effect of frequency of *following activity rules* behaviors at the conclusion of Week 2 may instead be indicative of an unaccounted factor: the unintended iatrogenic effects of grouping children with similarly deviant peers such that the delay discounting phenomenon is magnified (e.g., the smaller but sooner forms of social reinforcement from peers for off-task behavior is preferred over the larger but later forms of reinforcement provided by counselors) (M. E. Arnold & Hughes, 1999; Mellis et al., 2017). It is also noteworthy that children, irrespective of their concurrent pharmacotherapy status, achieved similar gains in *following activity rules, contributing to group discussion*, and *helping a peer* behaviors over time. This

provides further evidence for the effectiveness of the STP for children with and without prescribed medications at the outset of the intensive behavioral intervention.

Children did not demonstrate significant improvements in average number of negative behaviors across any behavior category or interval of time. These findings suggest that the response-cost system implemented in the 6-week STP is insufficient in regards to effective *punishment* of negative behaviors such as: *interrupting, whining, swearing, teasing, verbal abuse, leaving the activity area, intentional aggression*, and *intentional destruction*. When defined in terms of significant improvements in frequency of negative behaviors, there is no evidence in support of program effectiveness for the 6-week STP.

The 6-week STP did, however, yield significant improvements in children's academic and behavioral functioning in the academic learning center over time. The trend of improvement was significant for a quadratic relationship, such that there was a marked, linear increase in positive behaviors observed between Week 1 and Week 2, a more moderate linear increase in scores between Week 2 and Week 4, followed by a sharp decline in scores between Week 4 and Week 5. Generally, the ALC total scores trended upwards again from Week 5 to Week 6, although they did not reach the peak of positive behaviors observed at Week 4. Similar to the *following activity rules* behavior category, this quadratic trend of improvement in ALC total scores over time may be a reflection of the potentially iatrogenic effects of grouping deviant children together for behavioral treatment.

When measured in units of average percentage of idiosyncratic behavioral goals achieved each week, as measured by the DRC, the 6-week STP did not demonstrate significant

effects over time. However, findings suggest unmedicated and medicated children experience similar degrees of improvement in regards to individualized behavioral goals, with average percentage of DRC target behaviors achieved rising from 75.16% at Week 2 to 81.29% at Week 6. Given that children earned home-based privileges each day they earned ≥ 75% of their DRC goals, these findings yield clinically significant results: children, on average, gradually increased their daily percentage of DRC target behaviors achieved in a day over time. It is important to note that successful implementation of the DRC component of the STP requires caregivers' fidelity to the corresponding home-based contingency system. In the absence of caregivers' consistent and appropriate implementation of the home-based contingency system, the DRC intervention is effectively diluted. Similarly, if caregivers fail to offer sufficiently rewarding consequences for high DRC percentage scores and/or sufficiently aversive consequences for subpar DRC percentage scores, the potency of the DRC intervention is weakened. In the absence of measures of caregivers' fidelity to the DRC intervention, program effectiveness of the 6-week STP, in terms of percentage of DRC target behaviors achieved over time, is best described as indeterminate.

Moderators of Treatment Outcomes

The second aim of this study was to identify subpopulations of children with particularly strong or poor response to the 6-week STP. A significant strength of this study was the inclusion of children receiving concurrent pharmacotherapy, as this is a common clinical characteristic of children seeking treatment in community care settings. This also allowed for a comparison of behavioral outcomes among children based on their type(s) of concurrent

pharmacotherapy treatment. In the seminal MTA study, concurrent pharmacotherapy was limited to stimulants, with MPH being the most commonly prescribed medication treatment for study participants. In contrast, our sample lent itself to four discrete categories of medication treatment that children commonly present with in community care settings. However, study findings yielded nonsignificant moderating effects of concurrent pharmacotherapy status on behavioral outcomes over time. This suggests that children, irrespective of their concurrent pharmacotherapy status, experience similar responses to the 6-week STP intervention.

One of the two parent-reported measures of child functioning at pre-treatment, the BPM (covariate), demonstrated significant interactions with two categories of behavior defined by the STP point system: *following activity rules* and *swearing*. Secondary analyses of these significant interactions yielded a significant predictive relationship of the covariate on mean frequency counts of *following activity rules* behaviors at three time points. However, the predicted decrease in *following activity rules* frequency at Week 2 (-1.05), Week 3 (-1.14), and Week 4 (-1.13) did not communicate any clinically significant differences. The significant predictive relationship of the covariate on mean frequency counts of *swearing* behavior also proved to be clinically nonsignificant: at Week 5, a one point increase in the pre-treatment BPM total score predicted an increase in frequency of *swearing* behaviors at Week 5 (+.041).

Integration of Findings with Previous Research

Preliminary outcome studies of the standard STP model examined counselor reports of improvement and parent-ratings of ADHD, ODD, and CD symptom counts at post-treatment as measures of behavioral improvement and program effectiveness (Pelham & Hoza, 1996). Based

on counselor reports that indicated 91% of children were at least "somewhat improved" at post-treatment and parent-ratings which indicated statistically significant reductions in symptom counts for ADHD, ODD, and CD, the STP established robust acceptability and effectiveness. Since then, the STP has also demonstrated effectiveness in relation to social, recreational, and academic functioning (Chronis et al., 2004; Fabiano et al., 2007; Pelham et al., 2000, 2005, 2014, 2016). It is important to note, however, that racial minority and low-income families were underrepresented in these study samples. In the absence of sufficient sociodemographic diversity, the generalizability of these findings remains unclear. Adaptations of 2- and 3-week STP interventions have been reported to be effective treatments, referencing significant reductions in parent-reported symptoms of ADHD and ODD (Yamashita et al., 2010, 2011) and improvements in parent-reported measures of child functioning in relation to peer relationships and prosocial behaviors (Yamashita et al., 2010). Of note, these adapted STP interventions were developed and delivered in Japan; and further investigation is warranted to better understand cultural implications, if any, on treatment feasibility, acceptability, and effectiveness. Most recently, the Camp Baker pilot study (Tannenbaum et al., 2019) yielded statistically significant reductions in post-treatment parent-reported domains of child functional impairment, as demonstrated by lower symptom severity scores across externalizing, attention, and internalizing subscales on the BPM. The current study findings extend the existing literature by examining the effects of the 6-week STP intervention in terms of several functional behavioral outcome measures.

The first outcome measure was defined as parent-reported changes in child behaviors at post-treatment. The 6-week STP was associated with significant improvement on parent-reported measures of child functioning within the inattention, hyperactivity/impulsivity, and peer relations domains. It is important to note, however, that the parent-report behavior rating scale implemented in this study (Conners-3-P(S)) is not parallel to those used to operationalize child functioning in previous studies. As such, a direct comparison of treatment effectiveness between the adapted 6-week STP and prior iterations of the STP model cannot be completed at this time. As noted in previous studies, samples of children with ADHD are vastly heterogenous. In efforts to further investigate subpopulations that experience degrees of differential outcomes, it is common practice to categorize study participants into one or more subgroups of interest (e.g., gender, race/ethnicity, types of concurrent pharmacotherapy, pattern of comorbid diagnoses). Statistical analyses are then completed on both the total sample and the defined subgroup(s) to examine both treatment effects and moderators of outcomes. A notable limitation to this analytical approach is that by relying on a given research team's subjectively defined group(s) of interest, we are only able to estimate treatment effectiveness if children are able to neatly classify with one of the defined groups demonstrating significantly different responses to a treatment intervention. As such, there is a need to broaden the ways in which treatment effectiveness is defined and evaluated. Investigating individual rates of change, within a given study sample, serves to evaluate program effectiveness in terms of rates of clinically significant improvement or deterioration observed at post-treatment. This trend of analysis appears to be on the rise (Cuijpers et al., 2021; Döpfner et al., 2021; Murrell et al.,

2015; Siebelink et al., 2022), and was intentionally incorporated into this program effectiveness study to help establish one, more universal, comparison point for iterations of the STP model moving forward. The versatility of the reliable change (RC) analysis lends itself to any research study that measures an outcome variable using a validated measure with demonstrated test-retest reliability; and in this way a broader scope of interventions can be compared to each other in service of helping match families to the intervention that is best suited to their needs. In this study, the parent-report Conners-3-P(S) was administered at pre- and post-treatment. The established test-retest reliability and RCI values for each subscale (Conners, 2008) allowed for the RC analysis to be completed. Closely referencing the seminal work of Jacobson and Truax (1991), study participants' rates of reliable change (improvement or deterioration) among Conners-3-P(S) subscales were calculated. The absence of significantly greater rates of reliable improvement, as opposed to reliable deterioration, across *each* of the six domains of functioning on the Conners-3-P(S) was striking. Further still, when significantly greater rates of reliable improvement were indicated, less than half of participants attained reliable improvement (45%, hyperactivity/impulsivity; 37.5%, defiance/aggression) and, in one instance, the rate of reliable improvement was only marginally greater than the rate of reliable deterioration (37.5% v 32.5%, defiance/aggression). Curiously, patterns of significantly greater rates of reliable deterioration, as opposed to reliable improvement, were revealed within two concurrent pharmacotherapy groups: ADHD Only (38.9% v 33.3%, defiance/aggression) and ADHD + Other (60% v 40%, hyperactivity/impulsivity). It is important to note that the current study did not control for other variables that may have led to reduced rates of reliable

improvement, such as: child attendance, caregiver attendance to weekly behavioral parent training group, caregiver psychopathology at pre-treatment. Additionally, the outcome measures for the 6-week STP did not account for continuous evaluation of concurrent pharmacotherapy status, dosage(s), and medical adherence rates. To this end, these units of program effectiveness should be interpreted with caution.

The effectiveness of the 6-week STP was also evaluated in terms of: positive and negative behavior categories from the STP point system; classroom-based behaviors (i.e., ALC total score); and individualized treatment targets (e.g., DRC target behaviors). The 6-week STP yielded significant improvements in positive child behaviors (*following activity rules* and *helping a peer)* and classroom-based behaviors (e.g., seatwork completion and accuracy, appropriate behavior) comparable to the improvements demonstrated by the standard 8-week STP (Pelham et al., 2000). The 6-week STP also demonstrated nonsignificant effects on negative child behavior categories mirroring those nonsignificant effects of the standard STP. However, the 6-week STP did not yield significant improvements in DRC-based outcomes over time. This is in contrast to the standard 8-week STP which is associated with significant improvement in DRC-based outcomes $F(1, 99) = 5.94$; $p < .05$) and an average 10% increase in DRC scores from Week 2 to Week 8 (Pelham et al., 2000).

It is important to note that each STP outcome study to date has defined program objectives, and the corresponding measures of outcomes, in a multitude of ways. The current program evaluation was unique in that: the Conners-3-P(S) subscale *T*-scores at pre- and post-treatment were used as a proxy for parent-reported improvement in child behavior; the

negative behavior categories of *leaving activity area, intentional aggression, and intentional destruction* were evaluated as discrete categories (unlike previous studies where these behaviors were re-organized and labeled *conduct problems*); and behavioral functioning in the academic setting was evaluated as a single factor (unlike previous studies that have investigated discrete categories of behaviors in the classroom setting, such as: seatwork completion, accuracy, and productivity). Although this prevents direct comparisons of the effectiveness of the 6-week STP to other STP models at this time, the outcome measures used in this study were carefully considered and intentionally chosen to be measures that other STP programs, moving forward, can feasibly collect and analyze. This will then allow for a more accurate assessment of the 6-week STP's comparative effectiveness in terms of different behavioral outcome 'units'.

Limitations

Study Sample

Sample Size. A notable limitation of this study was the sample size. The participants included the 2019 cohort of Camp Baker which consisted of 52 children and their caregivers. Of these, only 48 children completed the 6-week STP. Due to the statistical influence of their outlying values, a further eight participants were omitted from one of the two primary analyses related to mean frequency counts of positive and negative behaviors defined by the STP point system. Although the realized sample sizes met or exceeded the sample sizes indicated by a priori power analysis as likely sufficient to detect small to moderate effects, the dichotomization of participants based on concurrent pharmacotherapy status led to an uneven number of children in each concurrent pharmacotherapy group (range: 5 – 18). This, in turn,

affected the power of primary statistical analyses that sought to examine moderating effects of treatment.

Assignment to Age-Matched Groups. Although the STP manual (Pelham et al., 2017; Synn et al., 2019) instructs that children should be organized into age-matched groups, group assignments are subjective and variable in that the Director of each discrete STP program is responsible for defining group criteria and applying it to variable samples of children presenting to treatment. Although not widely discussed, there is research to suggest that group treatments for externalizing behavior disorders may have unintended iatrogenic effects due to the way in which children seek out peer groups that affirm their maladaptive attitudes and behaviors (confluence model; Dishion et al., 1994). Over time, the strengthening of these peer relationships can magnify the delay aversion phenomenon children with ADHD experience and inadvertently undermine the effectiveness of a behavioral intervention (M. E. Arnold & Hughes, 1999; Dishion et al., 1999). To this end, the way in which groups are assigned and the impact these assignments may have on children's treatment outcomes warrants further examination and should be explored in future analyses.

Sociodemographic Diversity of Sample. The study sample was predominantly male (65.4%, $n = 46$) and Caucasian (65.4%, $n = 34$); and a majority of families reported a household income of at least $100,000 (73.1%, $n = 38$). To this end, the study sample was limited in terms of sociodemographic representativeness of the population at large. This, in turn, limits the generalizability of study findings. More importantly, the limited presence of racial minority and low-income families in research studies to date perpetuates, and in some cases exacerbates,

behavioral health disparities (Awad et al., 2022). The lack of sociodemographic diversity within this sample is symptomatic of our field's seemingly archaic recruitment, engagement, and research intervention strategies. Intentional research studies of program effectiveness across diverse sociodemographic samples is urgently needed to promote more equitable access to effective behavioral interventions.

Defined Subpopulations of Interest. Given the sometimes contradictory findings of which subpopulations of children have particularly strong or poor response to evidence-based interventions (Jensen et al., 2001; March et al., 2000; Owens et al., 2003) and the research-informed yet ultimately still subjective manner in which samples are organized into subgroups of interest, a more thoughtful and standardized approach to promote diversity within study samples is needed. Social determinants of health are factors of an individual's environment that confer greater risk for health disparities and poorer quality-of-life (US Department of Health and Services, 2022). Despite heightened nationwide social justice movements and the corresponding emphasis on the need to better recognize, understand, and advocate for diverse populations in an effort to reduce health disparities, *social determinants of health* surprisingly continue to be overlooked as a key factor of interest. Behavioral mental health providers and research would benefit from standardizing variables of social determinants of health (e.g., financial insecurity, neighborhood and built environment, housing instability, access to health care and/or preventative services, transportation difficulties, education access, and social and community context) as baseline characteristics to be accounted for and investigated as moderators of evidence-based treatment outcomes.

Imprecision of Measurement Protocols

Medication Status. In this study, children's concurrent pharmacotherapy status was established by referencing names of prescribed medications disclosed by parents at pre-treatment. This failed to take into account instances where children discontinued medication, were prescribed alternative medication(s) due to insufficient therapeutic effect, or demonstrated medical nonadherence (i.e., forgetting or refusing to take medication as prescribed). Future studies, particularly those investigating long-term outcomes, would benefit from inclusion of more frequent and detailed measures of medication status.

Behavior Rating Scales. The assessment battery specific to this study included two normed parent-report measures of child functioning, one of which was also used as a post-treatment outcome variable. To evaluate the extent to which each measure and/or subscale provided unique data about discrete constructs of clinical interest, correlations between all pre-treatment parent-reported variables of child functioning were calculated and reviewed. The correlations between the executive functioning, inattention, and hyperactivity/impulsivity subscales on the Conners-3-P(S) was as expected given previous research in support of the relationship between EF deficits and symptoms of inattention (Thorell, 2007; Willcutt et al., 2005) and the relationship between EF deficits and hyperactivity/impulsivity (Barkley, 1997; Sonuga-Barke, 2003). The three processes are believed to play uniquely significant roles in the dual pathway model of ADHD (Sonuga-Barke, 2002, 2003). However, the high degree of correlation between the executive functioning, inattention, and hyperactivity/impulsivity subscales maintains the possibility that these may not be three discrete constructs.

As is true of any rating scale, there is always the possibility of measurement error. RC analysis can help evaluate outcomes in a statistically reliable way that addresses the range in which scores may vary due to measurement error alone. Although this analysis was able to be used in this study, it was only feasible for one of the eighteen outcome measures. Thus, there is a strong likelihood that measurement error may have contributed to study findings, both statistically significant and nonsignificant.

Treatment Integrity and Fidelity. Although supervisory staff members observed groups regularly to monitor adherence to the treatment protocol and provide live-feedback as needed, Camp Baker did not implement structured assessments of treatment integrity and fidelity (e.g., treatment integrity and fidelity form; Chronis et al., 2004) or institute weekly point system reliability quizzes to help ensure staff members' accurate classification of behaviors in accordance with the STP treatment protocol. This study was also devoid of a measure of caregivers' fidelity to the home-based response cost system, that was an integral element of the DRC intervention. To this end, the reliability and validity of all behavioral observations remains undetermined; and study findings related to point system behaviors and DRC scores should be interpreted with caution.

ALC Point System. Previous studies have indicated that disruptive behaviors in the classroom, a common presenting concern for children with ADHD, can be a powerful antecedent to a decline in teachers' quality of classroom management behaviors (Williford & Vitiello, 2020). Moreover, preservice and early professional teachers have been reported to experience gaps in knowledge and implementation of classroom management strategies

(Poznanski et al., 2018). These findings suggest another possible factor in the fluctuating trends in average ALC scores over time. Although all Camp Baker staff participated in the same intensive training prior to the start of the program, it is possible that the training was insufficient for preservice teachers and teacher aides that served as ALC staff. Another consideration is the efficacy of the STP response-cost intervention in the academic learning center setting. Given the propensity for children with ADHD to demonstrate disruptive behaviors and/or poor academic performance, the ALC component of the STP may elicit a different category or intensity of negative behaviors over time such that the standard STP response-cost system does not effectively address the problem behaviors specific to this setting and its related tasks.

Statistical Analyses

A statistical shortcoming of this study was the sheer number of repeated measure ANCOVA analyses conducted, as this amount of testing results in inflated family-wide type I error. Although MANCOVA analyses would have been a more appropriate approach, the given study sample size did not yield sufficient power to conduct repeated measure MANCOVA analyses. The outcomes of interest in this study may have also been better served by an alternative generalized linear model approach: Poisson regression analysis (Agresti, 2015). This log-linear model is particularly appropriate for studies that have count data as the primary outcome (e.g., frequency of behavior). A Poisson regression analysis would also allow for determining the probability a certain number of child behaviors (e.g., positive, negative, ALC) occur over a given period of time (e.g., week; 6 weeks).

Implications

Although the STP is offered across fifteen sites in the US, there is a paucity of published research regarding the feasibility, acceptability, and effectiveness of the standard and adapted STP models across diverse sociodemographic populations. Fewer still are the number of studies completed without the involvement of one or more of the primary treatment developers. It can be argued that the research to date has significant methodological limitations and treatment providers should exercise caution when recommending the STP to children and their families until more rigorous program effectiveness studies are completed. Findings from this study give rise to important clinical and research considerations for providers and researchers moving forward.

Clinical

If parent-reported improvements in child functioning across inattention, hyperactivity/impulsivity, and peer relations domains or improvements in classroom-based behaviors are of primary interest, study findings indicate intensive behavioral treatment, such as the 6-week STP, may yield increments of desired change across one or more domains of child functioning. However, the scope of possible iatrogenic effects subsequent to group intervention remains unclear. In contrast, there continues to be strong evidence for the effectiveness of BPT interventions designed to provide caregivers with the level of psychoeducation and coaching needed for them to effectively utilize appropriate behavior management strategies themselves. What advantages one intervention uniquely offers, the other lacks. To balance the cost-effectiveness and feasibility of group interventions with the

robust efficacy of BPT interventions, it is possible a synthesis could be formed by creatively combining elements of both interventions and delivering them in the outpatient setting.

The principles of operant conditioning and social learning theory that underlie behavior management strategies for children with ADHD are not unlike those that underlie treatment engagement and skill mastery for caregivers of children with ADHD. In place of a counselor providing immediate and appropriate consequences for child behaviors (reinforcement for positive behaviors; punishment for negative behaviors) over 30 eight-hour camp days, perhaps a primary caregiver could receive an intensive training of their own (4- to 8-hours; one day) whereby therapists provide a high dose of psychoeducation, live-coaching, and opportunities to role-play particularly challenging parenting situations. Alternatively, or in addition to this approach, parent-child dyads could be served in an abbreviated intensive (e.g., 2 – 4 hours) pared down in scope while still allowing for live-coaching of caregivers and the opportunity to maximize child outcomes gleaned by the parallel process such a treatment model. In select cases, these approaches to shaping adaptive behaviors could also be abstracted to teachers and the school setting. If one or more of these supplemental interventions were applied to standard outpatient services or as an extension of the STP (i.e., offered in sequence to or contemporaneously with the STP), families would be one step closer to having increased access to the level of care the STP model provides without the time- and finance-based costs associated with the full STP model of care.

Further still, the STP and/or its elemental abstractions simultaneously serve as an invaluable training opportunity for unlicensed providers (e.g., preservice teachers,

undergraduate and graduate students, pre-doctoral interns, post-doctoral fellows). In this way, the evidence-based behavioral intervention can be provided at a lower-cost while granting licensed professionals in the outpatient setting greater bandwidth to serve families in need of services. This training model would likely lend itself to long-term sustainability, in terms of agencies' ability to staff and financially operate the STP intervention. This training model would also allow for expanded awareness among pre- and early-career psychologists, who may have had the opportunity to serve as a STP counselor, and these individuals' unique trajectories across the US would likely also contribute to greater dissemination and implementation of the STP and/or its adapted models.

Research

A notable strength of this program effectiveness study was the way in which it was designed to investigate clinically meaningful outcomes (i.e., rates of reliable change on measures of parent-reported child functioning at post-treatment) and socially valid target behaviors specific to children with ADHD, such as: frequency of prosocial versus invalid behaviors towards others as a measure of social functioning; rates of assignment completion and accuracy as a measure of academic functioning; and DRC achievement as a measure of children's acquisition of adaptive skill(s) pertaining to their individualized target behaviors. This study prompts further investigations of the moderators of treatment outcomes as a means of better understanding factors that contribute to children's differential rates of improvement over time. Over 50 years ago, Bergin and Strupp (1972) urged researchers to focus on differential outcomes between individuals, rather than simply investigating the efficacy of a given

intervention within a group sample, to better understand *for whom* and under *what circumstances* treatment is most effective. Single-case experimental design studies would allow for intersubject and intrasubject variability to be examined. This, in turn, would allow for more thorough investigation sof program effectiveness and help inform the development of more relevant adaptations to the STP model that better address moderating outcomes and allow for expanded access to high quality of care to children and families in need.

Prior to this study, two alternative adaptations to the standard STP had been developed and evaluated (Yamashita et al., 2010, 2011). Both the 2- and 3-week STP models yielded significant reductions in parent-reported symptoms of their child's ADHD and ODD at post-treatment. However, children who completed the 3-week STP did not demonstrate parent-reported improvements in peer relationships and prosocial behaviors (Yamashita et al., 2010). Of note, these studies implemented different measures of child functioning, despite a singular research team developing both the 2- and 3-week STP treatment models. One featured the Strengths and Difficulties Questionnaire (SDQ) (Yamashita et al., 2010) and the other reported outcomes obtained from validated measures of cognitive functioning (CogState; Health Solution, Inc.) (Yamashita et al., 2011). In the absence of corresponding RC analyses, a comprehensive evaluation of the adapted programs' clinical significance (i.e., rates of reliable improvement or deterioration, based on individual outcomes) remains undetermined. The RC analyses implemented in this study, to help evaluate program effectiveness in relation to parent-reported measures of child functioning at pre- and post-treatment, is a statistically sound approach that can be generalized to any program effectiveness study, regardless of the

specific measurement used (so long there is demonstrated evidence of its test-retest reliability). Although there is a lack of consensus as to what evidence-based assessment and monitoring tools should be established as the gold-standard for children with ADHD, more consistent implementation of RC analyses in outcome studies would allow for a meaningful step forward in the direction of improving benchmarks by which we compare the efficacy of two or more clinical interventions. This study has effectively established one value for this measure of clinically meaningful change (e.g., rates of reliable change at post-treatment using the Conners-3-P(S)). Additional values, generated by future STP evaluation studies, are needed to help providers and families more clearly understand the clinical implications of participating in one treatment model over another.

Lastly, given the evidence that ADHD, and disruptive behavior disorders, are chronic conditions, it would be prudent to evaluate the long-term impact of the STP and long-term models of care that incorporate the STP as one component (e.g., STP during the summer; School Consultation and/or Parent Training at scheduled intervals across the academic year; concurrent pharmacotherapy). Perhaps even more importantly, a multiple baseline design study would allow for discrete components of the STP intervention to be implemented sequentially and gradually, to control for threats to validity, while also allowing for more rigorous examination of how, and to what extent, each component of the STP intervention facilitates behavioral change.

Conclusions

This is the first study to investigate both program effectiveness and concurrent pharmacotherapy as a moderator of treatment outcome for child participants of a 6-week adaptation of the STP delivered in a community setting. Study findings supported the effectiveness of the adapted STP for children and families presenting with significant parent-reported concerns of child functioning across inattention, hyperactivity/impulsivity, and peer relations domains. Findings also indicated that children who complete the 6-week STP demonstrate significant improvements in adaptive behaviors, such as: *following activity rules, contributing to group discussion, and helping a peer*. There was also evidence to support the effectiveness of this 6-week STP in improving children's behavioral and academic functioning in a setting that mirrors the classroom environment. However, when behavioral improvement was measured in terms of parent-reported concerns of child functioning across learning problems, executive functioning, and defiance/aggression domains; frequency of positive point system behaviors (e.g., *answering attention questions, complying with a command*, and *ignoring a negative stimulus*); frequency of negative point system behaviors (e.g., *interrupting, whining, swearing, teasing, verbal abuse, leaving the activity area, intentional aggression*, and *intentional destruction*); and percentage of DRC target behaviors achieved over time, the 6-week STP yielded nonsignificant findings. Furthermore, there was no evidence to support the moderating effect of concurrent pharmacotherapy type on any of the aforementioned behavioral outcomes.

In the absence of clear findings to guide treatment recommendations and families' informed decision to participate in treatment, further investigation of the both the therapeutic and adverse outcomes of the STP models is urgently needed. This study provides a standard for

behavioral outcome measures that can be replicated in future STP research study protocols, providing a clearer point of comparison between STP teams' and, most notably, communicate rates of parent-reported improvement and/or deterioration across domains of functioning that are particularly meaningful to caregivers as they consider the costs and benefits of participating in one treatment intervention over another.

Appendix A: List of STP Point System Behavior Categories

Positive Behavior Categories

Positive Interval Categories (points awarded at Point Check [a])

1. Following Activity Rules	+ 50 points per interval
2. Good Sportsmanship	+ 25 points per interval
3. Behavior Bonus	+ 25 points per interval

Positive Frequency Categories (points awarded when behavior occurs)

4. Standardized Attention	+ 10 points
5. Non-standardized Attention	+ 10 points
6. Compliance	+ 10 points
7. Helping a Peer	+ 10 points
8. Sharing with a Peer	+ 10 points
9. Contributing to a Group Discussion	+ 10 points
10. Ignoring a Negative Stimulus	+ 25 points

Negative Behavior Categories (points deducted when behavior occurs)

1. Violating Activity Rules	- 10 points
2. Poor Sportsmanship	- 10 points

Negative Physical Categories

3. Intentional Aggression (toward a peer or staff member)	- 50 points
4. Unintentional Aggression (toward a peer or staff member)	- 50 points
5. Intentional Destruction of Property	- 50 points
6. Unintentional Destruction of Property	- 50 points
7. Noncompliance / Repeated Noncompliance	- 20 points
8. Stealing	- 50 points
9. Leaving the Activity Area Without Permission	- 50 points

Negative Verbal Categories

10. Lying	- 20 points
11. Verbal Abuse to Staff	- 20 points
12. Name Calling / Teasing	- 20 points
13. Cursing / Swearing	- 20 points
14. Interruption	- 20 points
15. Complaining / Whining	- 20 points

[a] Children earn points for each fixed interval during which they do not violate criteria; points are awarded at point checks conducted at the end of each activity.

Appendix B: Example Daily Report Card

Child Name: _____ Date: _____

Target Behavior	AM Interval	PM Interval	Met Target in AM?	Met Target in PM?
Earn Following Activity Rules for 2 out of 4 intervals in AM and PM			Y N NA	Y N NA
Contribute to Group Discussion 1 time in AM and 1 time in PM			Y N NA	Y N NA
No more than 5 rule violations (tallies) in Academics			Y N NA	Y N NA
Ask 1 new friend 1 question in AM and 1 question in PM			Y N NA	Y N NA

Earned Lunchtime Recess? …….. No Yes
Earned End of Day Recess? …….. No Yes

Overall Number of "Yes" Marks _____ Overall Number of "No" Marks _____

Percentage Positive _____

Comments:

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis: Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459. https://doi.org/10.1002/wics.101

Abikoff, H., Courtney, M., Pelham, W. E., & Koplewicz, H. S. (1993). Teachers' ratings of disruptive behaviors: The influence of halo effects. *Journal of Abnormal Child Psychology*, *21*(5), 519–533. https://doi.org/10.1007/BF00916317

Abikoff, H., Hechtman, L., Klein, R. G., Gallagher, R., Fleiss, K., Etcovitch, J., Cousins, L., Greenfield, B., Martin, D., & Pollack, S. (2004). Social Functioning in Children With ADHD Treated With Long-Term Methylphenidate and Multimodal Psychosocial Treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(7), 820–829. https://doi.org/10.1097/01.chi.0000128797.91601.1a

Achenbach, T. M., Ivanova, M. Y., & Rescorla, L. A. (2019). Multi-Informant Assessment of Psychopathology from Preschool through Old Age. In M. Sellbom & J. A. Suhr (Eds.), *The Cambridge Handbook of Clinical Assessment and Diagnosis* (1st ed., pp. 123–134). Cambridge University Press. https://doi.org/10.1017/9781108235433.011

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213–232. https://doi.org/10.1037/0033-2909.101.2.213

Achenbach, T., McConaughy, S., Ivanova, M., & Rescorla, L. (2011). *Manual for the ASEBA Brief Problem Monitor (BPM)*. Burlington: University of Vermont, Research Center for Children, Youth, & Families. https://documents.acer.org/ASEBA_Brief_Problem_Monitor_Manual.pdf

Achenbach, T., & Rescorla, L. (2001). *Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Multi-Informant Assessment.* Burlington: University of Vermont, Research Center for Children, Youth, & Families.

Agresti, A. (2015). *Foundations of linear and generalized linear models*. John Wiley & Sons Inc.

Albajara Sáenz, A., Villemonteix, T., & Massat, I. (2019). Structural and functional neuroimaging in attention-deficit/hyperactivity disorder. *Developmental Medicine & Child Neurology*, *61*(4), 399–405. https://doi.org/10.1111/dmcn.14050

American Academy of Pediatrics. Subcommittee on Attention-Deficit/Hyperactivity Disorder and Committee on Quality Improvement. (2001). Clinical practice guideline: Treatment of the school-aged child with attention-deficit/hyperactivity disorder. *Pediatrics*, *108*(4), 1033–1044. https://doi.org/10.1542/peds.108.4.1033

American Psychiatric Association (Ed.). (1968). *Diagnostic and statistical manual of mental disorders: DSM-II* (2. ed). American Psychiatric Publishing.

American Psychiatric Association (Ed.). (1980). *Diagnostic and statistical manual of mental disorders: DSM-III* (3. ed). American Psychiatric Publishing.

American Psychiatric Association (Ed.). (1987). *Diagnostic and statistical manual of mental disorders: DSM-III-R* (3. ed, Revised). American Psychiatric Publishing.

American Psychiatric Association (Ed.). (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV* (4. ed). American Psychiatric Publishing.

American Psychiatric Association (Ed.). (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5. ed). American Psychiatric Publishing.

Arnold, L. E., Abikoff, H. B., Cantwell, D. P., Conners, C. K., Elliott, G., Greenhill, L. L., Hechtman, L., Hinshaw, S. P., Hoza, B., Jensen, P. S., Kraemer, H. C., March, J. S., Newcorn, J. H., Pelham, W. E., Richters, J. E., Schiller, E., Severe, J. B., Swanson, J. M., Vereen, D., & Wells, K. C. (1997). National Institute of Mental Health Collaborative Multimodal Treatment Study of Children with ADHD (the MTA). Design challenges and choices. *Archives of General Psychiatry*, *54*(9), 865–870. https://doi.org/10.1001/archpsyc.1997.01830210113015

Arnold, L. E., Chuang, S., Davies, M., Abikoff, H. B., Conners, C. K., Elliott, G. R., Greenhill, L. L., Hechtman, L., Hinshaw, S. P., Hoza, B., Jensen, P. S., Kraemer, H. C., Langworthy-Lam, K. S., March, J. S., Newcorn, J. H., Pelham, W. E., Severe, J. B., Swanson, J. M., Vitiello, B., … Wigal, T. (2004). Nine months of multicomponent behavioral treatment for ADHD and effectiveness of MTA fading procedures. *Journal of Abnormal Child Psychology*, *32*(1), 39–51. https://doi.org/10.1023/b:jacp.0000007579.61289.31

Arnold, M. E., & Hughes, J. N. (1999). First Do No Harm. *Journal of School Psychology*, *37*(1), 99–115. https://doi.org/10.1016/S0022-4405(98)00027-2

Arnsten, A. F. T. (2006). Stimulants: Therapeutic Actions in ADHD. *Neuropsychopharmacology*, *31*(11), 2376–2383. https://doi.org/10.1038/sj.npp.1301164

Awad, G. H., Abuelezam, N. N., Ajrouch, K. J., & Stiffler, M. J. (2022). Lack of Arab or Middle Eastern and North African Health Data Undermines Assessment of Health Disparities. *American Journal of Public Health*, *112*(2), 209–212. https://doi.org/10.2105/AJPH.2021.306590

Bandura, A. (1977). *Social learning theory.* Prentice-Hall.

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2007a). Long-Term School Outcomes for Children with Attention-Deficit/Hyperactivity Disorder: A Population-Based Perspective: *Journal of Developmental & Behavioral Pediatrics*, *28*(4), 265–273. https://doi.org/10.1097/DBP.0b013e31811ff87d

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2007b). Modifiers of Long-Term School Outcomes for Children with Attention-Deficit/Hyperactivity Disorder: Does Treatment with Stimulant Medication Make a Difference? Results from a Population-Based Study: *Journal of Developmental & Behavioral Pediatrics*, *28*(4), 274–287. https://doi.org/10.1097/DBP.0b013e3180cabc28

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94. https://doi.org/10.1037/0033-2909.121.1.65

Barkley, R. A. (Ed.). (2015). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (Fourth edition). The Guilford Press.

Barkley, R. A., & Fischer, M. (2019). Hyperactive Child Syndrome and Estimated Life Expectancy at Young Adult Follow-Up: The Role of ADHD Persistence and Other Potential Predictors.

Journal of Attention Disorders, *23*(9), 907–923. https://doi.org/10.1177/1087054718816164

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2006). Young Adult Outcome of Hyperactive Children: Adaptive Functioning in Major Life Activities. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(2), 192–202. https://doi.org/10.1097/01.chi.0000189134.97436.e2

Barry, C. T., Lindsey, R. A., & Neumann, A. A. (2019). Assessment of Childhood Disruptive Behavior Disorders and Attention-Deficit/Hyperactivity Disorder. In M. Sellbom & J. A. Suhr (Eds.), *The Cambridge Handbook of Clinical Assessment and Diagnosis* (1st ed., pp. 308–316). Cambridge University Press. https://doi.org/10.1017/9781108235433.022

Barry, T. D., Lyman, R. D., & Klinger, L. G. (2002). Academic Underachievement and Attention-Deficit/Hyperactivity Disorder. *Journal of School Psychology*, *40*(3), 259–283. https://doi.org/10.1016/S0022-4405(02)00100-0

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs*, *75*(1), 43–88.

Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology*, *4*(1, Pt.2), 1–103. https://doi.org/10.1037/h0030372

Bergin, A. E., & Strupp, H. H. (1972). *Changing frontiers in the science of psychotherapy*. Aldine · Atherton.

Biederman, J. (1996). A Prospective 4-Year Follow-up Study of Attention-Deficit Hyperactivity and Related Disorders. *Archives of General Psychiatry*, *53*(5), 437. https://doi.org/10.1001/archpsyc.1996.01830050073012

Biederman, J., Faraone, S., Milberger, S., Curtis, S., Chen, L., Marrs, A., Ouellette, C., Moore, P., & Spencer, T. (1996). Predictors of persistence and remission of ADHD into adolescence: Results from a four-year prospective follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(3), 343–351. https://doi.org/10.1097/00004583-199603000-00016

Biederman, J., Kwon, A., Aleardi, M., Chouinard, V.-A., Marino, T., Cole, H., Mick, E., & Faraone, S. V. (2005). Absence of Gender Effects on Attention Deficit Hyperactivity Disorder: Findings in Nonreferred Subjects. *American Journal of Psychiatry*, *162*(6), 1083–1089. https://doi.org/10.1176/appi.ajp.162.6.1083

Biederman, J., Mick, E., & Faraone, S. V. (1998). Depression in attention deficit hyperactivity disorder (ADHD) children: "True" depression or demoralization? *Journal of Affective Disorders*, *47*(1–3), 113–122. https://doi.org/10.1016/S0165-0327(97)00127-4

Biederman, J., Monuteaux, M. C., Spencer, T., Wilens, T. E., & Faraone, S. V. (2009). Do Stimulants Protect Against Psychiatric Disorders in Youth With ADHD? A 10-Year Follow-up Study. *PEDIATRICS*, *124*(1), 71–78. https://doi.org/10.1542/peds.2008-3347

Biederman, J., Petty, C. R., Clarke, A., Lomedico, A., & Faraone, S. V. (2011). Predictors of persistent ADHD: An 11-year follow-up study. *Journal of Psychiatric Research*, *45*(2), 150–155. https://doi.org/10.1016/j.jpsychires.2010.06.009

Brown, R. T., Antonuccio, D. O., DuPaul, G. J., Fristad, M. A., King, C. A., Leslie, L. K., McCormick, G. S., Pelham, W. E., Piacentini, J. C., & Vitiello, B. (2008). Attention-deficit/hyperactivity disorder. In R. T. Brown, D. O. Antonuccio, G. J. DuPaul, M. A. Fristad, C. A. King, L. K. Leslie, G. S. McCormick, W. E. Pelham, J. C. Piacentini, & B. Vitiello, *Childhood mental health disorders: Evidence base and contextual factors for psychosocial, psychopharmacological, and combined interventions.* (pp. 15–32). American Psychological Association. https://doi.org/10.1037/11638-002

Cantwell, D. P. (1996). Attention Deficit Disorder: A Review of the Past 10 Years. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(8), 978–987. https://doi.org/10.1097/00004583-199608000-00008

Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends in Cognitive Sciences*, *10*(3), 117–123. https://doi.org/10.1016/j.tics.2006.01.011

Cervantes, P. E., Matson, J. L., Adams, H. L., Williams, L. W., Goldin, R. L., & Jang, J. (2013). Comparing social skill profiles of children with autism spectrum disorders versus children with attention deficit hyperactivity disorder: Where the deficits lie. *Research in Autism Spectrum Disorders*, *7*(9), 1104–1110. https://doi.org/10.1016/j.rasd.2013.05.008

Chang, Z., D'Onofrio, B. M., Quinn, P. D., Lichtenstein, P., & Larsson, H. (2016). Medication for Attention-Deficit/Hyperactivity Disorder and Risk for Depression: A Nationwide

Longitudinal Cohort Study. *Biological Psychiatry*, *80*(12), 916–922. https://doi.org/10.1016/j.biopsych.2016.02.018

Chang, Z., Lichtenstein, P., Halldner, L., D'Onofrio, B., Serlachius, E., Fazel, S., Långström, N., & Larsson, H. (2014). Stimulant ADHD medication and risk for substance abuse. *Journal of Child Psychology and Psychiatry*, *55*(8), 878–885. https://doi.org/10.1111/jcpp.12164

Chen, W. J., Faraone, S. V., Biederman, J., & Tsuang, M. T. (1994). Diagnostic accuracy of the Child Behavior Checklist scales for attention-deficit hyperactivity disorder: A receiver-operating characteristic analysis. *Journal of Consulting and Clinical Psychology*, *62*(5), 1017–1025. https://doi.org/10.1037/0022-006X.62.5.1017

Chronis, A. M., Chacko, A., Fabiano, G. A., Wymbs, B. T., & Pelham, Jr., W. E. (2004). Enhancements to the Behavioral Parent Training Paradigm for Families of Children with ADHD: Review and Future Directions. *Clinical Child and Family Psychology Review*, *7*(1), 1–27. https://doi.org/10.1023/B:CCFP.0000020190.60808.a4

Chronis, A. M., Jones, H. A., & Raggi, V. L. (2006). Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. *Clinical Psychology Review*, *26*(4), 486–502. https://doi.org/10.1016/j.cpr.2006.01.002

Chronis-Tuscano, A., O'Brien, K. A., Johnston, C., Jones, H. A., Clarke, T. L., Raggi, V. L., Rooney, M. E., Diaz, Y., Pian, J., & Seymour, K. E. (2011). The Relation Between Maternal ADHD Symptoms & Improvement in Child Behavior Following Brief Behavioral Parent Training is Mediated by Change in Negative Parenting. *Journal of Abnormal Child Psychology*, *39*(7), 1047–1057. https://doi.org/10.1007/s10802-011-9518-2

Chronis-Tuscano, A., Raggi, V. L., Clarke, T. L., Rooney, M. E., Diaz, Y., & Pian, J. (2008). Associations between Maternal Attention-Deficit/Hyperactivity Disorder Symptoms and Parenting. *Journal of Abnormal Child Psychology*, *36*(8), 1237–1250. https://doi.org/10.1007/s10802-008-9246-4

Coker, T. R., Elliott, M. N., Toomey, S. L., Schwebel, D. C., Cuccaro, P., Tortolero Emery, S., Davies, S. L., Visser, S. N., & Schuster, M. A. (2016). Racial and Ethnic Disparities in ADHD Diagnosis and Treatment. *PEDIATRICS*, *138*(3), e20160407–e20160407. https://doi.org/10.1542/peds.2016-0407

Coles, E. K., Pelham III, W. E., Fabiano, G. A., Gnagy, E. M., Burrows-MacLean, L., Wymbs, B. T., Chacko, A., Walker, K. S., Wymbs, F., Robb Mazzant, J., Garefino, A., Hoffman, M. T., Massetti, G. M., Page, T. F., Waschbusch, D. A., Waxmonsky, J. G., & Pelham Jr., W. E. (2019). Randomized Trial of First-Line Behavioral Intervention to Reduce Need for Medication in Children with ADHD. *Journal of Clinical Child & Adolescent Psychology*, 1–15. https://doi.org/10.1080/15374416.2019.1630835

Coles, E. K., Pelham, W. E., Gnagy, E. M., Burrows-Maclean, L., Fabiano, G. A., Chacko, A., Wymbs, B. T., Tresco, K. E., Walker, K. S., & Robb, J. A. (2005). A Controlled Evaluation of Behavioral Treatment With Children With ADHD Attending a Summer Treatment Program. *Journal of Emotional and Behavioral Disorders*, *13*(2), 99–112. https://doi.org/10.1177/10634266050130020301

Conners, C. (2008). *Conners 3rd Edition Manual*. Toronto: Multi-Health Systems.

Cortese, S., Ramos Olazagasti, M. A., Klein, R. G., Castellanos, F. X., Proal, E., & Mannuzza, S. (2013). Obesity in Men With Childhood ADHD: A 33-Year Controlled, Prospective, Follow-up Study. *PEDIATRICS*, *131*(6), e1731–e1738. https://doi.org/10.1542/peds.2012-0540

Cuffe, S. P., Visser, S. N., Holbrook, J. R., Danielson, M. L., Geryk, L. L., Wolraich, M. L., & McKeown, R. E. (2015). ADHD and Psychiatric Comorbidity: Functional Outcomes in a School-Based Sample of Children. *Journal of Attention Disorders*, 108705471561343. https://doi.org/10.1177/1087054715613437

Cuijpers, P., Karyotaki, E., Ciharova, M., Miguel, C., Noma, H., Stikkelbroek, Y., Weisz, J. R., & Furukawa, T. A. (2021). The effects of psychological treatments of depression in children and adolescents on response, reliable change, and deterioration: A systematic review and meta-analysis. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-021-01884-6

Dalsgaard, S., Østergaard, S. D., Leckman, J. F., Mortensen, P. B., & Pedersen, M. G. (2015). Mortality in children, adolescents, and adults with attention deficit hyperactivity disorder: A nationwide cohort study. *The Lancet*, *385*(9983), 2190–2196. https://doi.org/10.1016/S0140-6736(14)61684-6

Danielson, M. L., Bitsko, R. H., Ghandour, R. M., Holbrook, J. R., Kogan, M. D., & Blumberg, S. J. (2018). Prevalence of Parent-Reported ADHD Diagnosis and Associated Treatment Among U.S. Children and Adolescents, 2016. *Journal of Clinical Child & Adolescent Psychology*, *47*(2), 199–212. https://doi.org/10.1080/15374416.2017.1417860

Daviss, W. B. (2008). A Review of Co-Morbid Depression in Pediatric ADHD: Etiologies, Phenomenology, and Treatment. *Journal of Child and Adolescent Psychopharmacology*, *18*(6), 565–571. https://doi.org/10.1089/cap.2008.032

Deault, L. C. (2010). A Systematic Review of Parenting in Relation to the Development of Comorbidities and Functional Impairments in Children with Attention-Deficit/Hyperactivity Disorder (ADHD). *Child Psychiatry & Human Development*, *41*(2), 168–192. https://doi.org/10.1007/s10578-009-0159-4

Deniston, O. L., Rosenstock, I. M., & Getting, V. A. (1968). *Evaluation of Program Effectiveness.* *83*(4), 323–335.

Diamantopoulou, S., Rydell, A.-M., Thorell, L. B., & Bohlin, G. (2007). Impact of Executive Functioning and Symptoms of Attention Deficit Hyperactivity Disorder on Children's Peer Relations and School Performance. *Developmental Neuropsychology*, *32*(1), 521–542. https://doi.org/10.1080/87565640701360981

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm. Peer groups and problem behavior. *The American Psychologist*, *54*(9), 755–764. https://doi.org/10.1037//0003-066x.54.9.755

Dishion, T. J., Patterson, G. R., & Griesler, P. C. (1994). Peer Adaptations in the Development of Antisocial Behavior. In L. R. Huesmann (Ed.), *Aggressive Behavior* (pp. 61–95). Springer US. https://doi.org/10.1007/978-1-4757-9116-7_4

Döpfner, M., Liebermann-Jordanidis, H., Kinnen, C., Hallberg, N., Mokros, L., Benien, N., Mütsch, A., Schürmann, S., Wolff Metternich-Kaizman, T., Hautmann, C., & Dose, C.

(2021). Long-Term Effectiveness of Guided Self-Help for Parents of Children With ADHD in Routine Care-An Observational Study. *Journal of Attention Disorders*, *25*(2), 265–274. https://doi.org/10.1177/1087054718810797

Douglas, V. I. (1972). Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *4*(4), 259–282. https://doi.org/10.1037/h0082313

Dupaul, G. J., Mcgoey, K. E., Eckert, T. L., & Vanbrakle, J. (2001). Preschool Children With Attention-Deficit/Hyperactivity Disorder: Impairments in Behavioral, Social, and School Functioning. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(5), 508–515. https://doi.org/10.1097/00004583-200105000-00009

Elia, J., Ambrosini, P., & Berrettini, W. (2008). ADHD characteristics: I. Concurrent co-morbidity patterns in children & adolescents. *Child and Adolescent Psychiatry and Mental Health*, *2*(1), 15. https://doi.org/10.1186/1753-2000-2-15

Evans, S. W., Owens, J. S., Wymbs, B. T., & Ray, A. R. (2018). Evidence-Based Psychosocial Treatments for Children and Adolescents With Attention Deficit/Hyperactivity Disorder. *Journal of Clinical Child & Adolescent Psychology*, *47*(2), 157–198. https://doi.org/10.1080/15374416.2017.1390757

Eyberg, S. (2005). Tailoring and adapting parent-child interaction therapy to new populations. *Education and Treatment of Children*, *28*(2), 197–201.

Fabiano, G. A., Schatz, N. K., & Pelham, W. E. (2014). Summer treatment programs for youth with ADHD. *Child and Adolescent Psychiatric Clinics of North America*, *23*(4), 757–773. https://doi.org/10.1016/j.chc.2014.05.012

Fabiano, G. A., Vujnovic, R. K., Pelham, W. E., Waschbusch, D. A., Massetti, G. M., Pariseau, M. E., Naylor, J., Yu, J., Robins, M., Carnefix, T., Greiner, A. R., & Volker, M. (2010). Enhancing the Effectiveness of Special Education Programming for Children With Attention Deficit Hyperactivity Disorder Using a Daily Report Card. *School Psychology Review*, *39*(2), 219–239. https://doi.org/10.1080/02796015.2010.12087775

Faraone, S. V. (2015). Attention deficit hyperactivity disorder and premature death. *The Lancet*, *385*(9983), 2132–2133. https://doi.org/10.1016/S0140-6736(14)61822-5

Faraone, S. V., & Wilens, T. E. (2007). Effect of stimulant medications for attention-deficit/hyperactivity disorder on later substance use and the potential for stimulant misuse, abuse, and diversion. *The Journal of Clinical Psychiatry*, *68 Suppl 11*, 15–22.

Farmer, R. F., & Chapman, A. L. (2016). Changing behavior by changing the environment. In R. F. Farmer & A. L. Chapman, *Behavioral interventions in cognitive behavior therapy: Practical guidance for putting theory into action (2nd ed.).* (pp. 101–132). American Psychological Association. https://doi.org/10.1037/14691-004

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Felt, B. T., Christner, J. G., Kochhar, P., & Van Harrison, R. (2014). Diagnosis and Management of ADHD in Children. *Am Fam Physician*, *90*(7), 456–464.

Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting and Clinical Psychology*, *67*(3), 308–319. https://doi.org/10.1037/0022-006X.67.3.308

Frazier, T. W., Demaree, H. A., & Youngstrom, E. A. (2004). Meta-Analysis of Intellectual and Neuropsychological Test Performance in Attention-Deficit/Hyperactivity Disorder. *Neuropsychology*, *18*(3), 543–555. https://doi.org/10.1037/0894-4105.18.3.543

Frazier, T. W., Youngstrom, E. A., Glutting, J. J., & Watkins, M. W. (2007). ADHD and Achievement: Meta-Analysis of the Child, Adolescent, and Adult Literatures and a Concomitant Study With College Students. *Journal of Learning Disabilities*, *40*(1), 49–65. https://doi.org/10.1177/00222194070400010401

Froehlich, T. E., Lanphear, B. P., Epstein, J. N., Barbaresi, W. J., Katusic, S. K., & Kahn, R. S. (2007). Prevalence, Recognition, and Treatment of Attention-Deficit/Hyperactivity Disorder in a National Sample of US Children. *Archives of Pediatrics & Adolescent Medicine*, *161*(9), 857. https://doi.org/10.1001/archpedi.161.9.857

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. *Statistical Science*, *24*(3). https://doi.org/10.1214/09-STS301

Gaub, M., & Carlson, C. L. (1997). Gender Differences in ADHD: A Meta-Analysis and Critical Review. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(8), 1036–1045. https://doi.org/10.1097/00004583-199708000-00011

Gerdes, A. C., Hoza, B., Arnold, L. E., Pelham, W. E., Swanson, J. M., Wigal, T., & Jensen, P. S. (2007). Maternal depressive symptomatology and parenting behavior: Exploration of possible mediators. *Journal of Abnormal Child Psychology*, *35*(5), 705–714. https://doi.org/10.1007/s10802-007-9134-3

Gershon, J. (2002). A Meta-Analytic Review of Gender Differences in ADHD. *Journal of Attention Disorders*, *5*(3), 143–154. https://doi.org/10.1177/108705470200500302

Gomez, R., & Vance, A. (2019). Conners 3-Parent (Short): Measurement Invariance Across Gender, Concurrent and Discriminant Validities. *Journal of Childhood & Developmental Disorders*, *05*(01). https://doi.org/10.4172/2472-1786.100079

Harvey, E. A., Breaux, R. P., & Lugo-Candelas, C. I. (2016). Early development of comorbidity between symptoms of attention-deficit/hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *Journal of Abnormal Psychology*, *125*(2), 154–167. https://doi.org/10.1037/abn0000090

Harvey, E., Danforth, J. S., McKee, T. E., Ulaszek, W. R., & Friedman, J. L. (2003). Parenting of children with Attention-Defecit/Hyperactivity Disorder (ADHD): The role of parental ADHD symptomatology. *Journal of Attention Disorders*, *7*(1), 31–42. https://doi.org/10.1177/108705470300700104

Hechtman, L., Swanson, J. M., Sibley, M. H., Stehli, A., Owens, E. B., Mitchell, J. T., Arnold, L. E., Molina, B. S. G., Hinshaw, S. P., Jensen, P. S., Abikoff, H. B., Perez Algorta, G., Howard, A. L., Hoza, B., Etcovitch, J., Houssais, S., Lakes, K. D., Nichols, J. Q., Vitiello, B., … Stern, K. (2016). Functional Adult Outcomes 16 Years After Childhood Diagnosis of Attention-Deficit/Hyperactivity Disorder: MTA Results. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(11), 945–952.e2. https://doi.org/10.1016/j.jaac.2016.07.774

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. JSTOR.

Hoza, B., Mrug, S., Gerdes, A. C., Hinshaw, S. P., Bukowski, W. M., Gold, J. A., Kraemer, H. C., Pelham, W. E., Wigal, T., & Arnold, L. E. (2005). What Aspects of Peer Relationships Are Impaired in Children With Attention-Deficit/Hyperactivity Disorder? *Journal of Consulting and Clinical Psychology*, *73*(3), 411–423. https://doi.org/10.1037/0022-006X.73.3.411

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037//0022-006x.59.1.12

Jensen, P. S., Hinshaw, S. P., Kraemer, H. C., Lenora, N., Newcorn, J. H., Abikoff, H. B., March, J. S., Arnold, L. E., Cantwell, D. P., Conners, C. K., Elliott, G. R., Greenhill, L. L., Hechtman, L., Hoza, B., Pelham, W. E., Severe, J. B., Swanson, J. M., Wells, K. C., Wigal, T., & Vitiello, B. (2001). ADHD comorbidity findings from the MTA study: Comparing comorbid

subgroups. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(2), 147–158. https://doi.org/10.1097/00004583-200102000-00009

Johnston, C., & Murray, C. (2003). Incremental Validity in the Psychological Assessment of Children and Adolescents. *Psychological Assessment*, *15*(4), 496–507. https://doi.org/10.1037/1040-3590.15.4.496

Jones, H. A., Epstein, J. N., Hinshaw, S. P., Owens, E. B., Chi, T. C., Arnold, L. E., Hoza, B., & Wells, K. C. (2010). Ethnicity as a Moderator of Treatment Effects on Parent—Child Interaction for Children With ADHD. *Journal of Attention Disorders*, *13*(6), 592–600. https://doi.org/10.1177/1087054709332158

Kazdin, A. E. (1997). Parent Management Training: Evidence, Outcomes, and Issues. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(10), 1349–1356. https://doi.org/10.1097/00004583-199710000-00016

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, *63*(3), 146–159. https://doi.org/10.1037/0003-066X.63.3.146

Kazdin, A. E. (2009). *Parent management training: Treatment for oppositional, aggressive, and antisocial behavior in children and adolescents*. Oxford University Press.

Kelley, M. L., & McCain, A. P. (1995). Promoting Academic Performance in Inattentive Children: The Relative Efficacy of School-Home Notes with and without Response Cost. *Behavior Modification*, *19*(3), 357–375. https://doi.org/10.1177/01454455950193006

Kirk, S. A. (2004). Are Children's DSM Diagnoses Accurate? *Brief Treatment and Crisis Intervention*, *4*(3), 255–270. https://doi.org/10.1093/brief-treatment/mhh022

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., Barkley, R. A., Newcorn, J., Jensen, P., & Richters, J. (1994). DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *The American Journal of Psychiatry*, *151*(11), 1673–1685. https://doi.org/10.1176/ajp.151.11.1673

Lahey, B. B., Lee, S. S., Sibley, M. H., Applegate, B., Molina, B. S. G., & Pelham, W. E. (2016). Predictors of adolescent outcomes among 4–6–year–old children with attention–deficit/hyperactivity disorder. *Journal of Abnormal Psychology*, *125*(2), 168–181. https://doi.org/10.1037/abn0000086

Lahey, B. B., Pelham, W. E., Stein, M. A., Loney, J., Trapani, C., Nugent, K., Kipp, H., Schmidt, E., Lee, S., Cale, M., Gold, E., Hartung, C. M., Willcutt, E., & Baumann, B. (1998). Validity of DSM-IV Attention-Deficit/Hyperactivity Disorder for Younger Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *37*(7), 695–702. https://doi.org/10.1097/00004583-199807000-00008

Landrum, T. J., & Kauffman, J. M. (2006). Behavioral approaches to classroom management. In *Handbook of classroom management: Research, practice, and contemporary issues.* (pp. 47–71).

Lange, K. W., Reichl, S., Lange, K. M., Tucha, L., & Tucha, O. (2010). The history of attention deficit hyperactivity disorder. *ADHD Attention Deficit and Hyperactivity Disorders*, *2*(4), 241–255. https://doi.org/10.1007/s12402-010-0045-8

LeFever, G. B., Villers, M. S., Morrow, A. L., & Vaughn, E. S. (2002). Parental perceptions of adverse educational outcomes among children diagnosed and treated for ADHD: A call for improved school/provider collaboration. *Psychology in the Schools*, *39*(1), 63–71. https://doi.org/10.1002/pits.10000

Lewis-Fernández, R., & Díaz, N. (2002). The Cultural Formation: A Method for Assessing Cultural Factors Affecting the Clinical Encounter. *Psychiatric Quarterly*, *73*(4), 271–295. https://doi.org/10.1023/A:1020412000183

Liang, J., Matheson, B. E., & Douglas, J. M. (2016). Mental Health Diagnostic Considerations in Racial/Ethnic Minority Youth. *Journal of Child and Family Studies*, *25*(6), 1926–1940. https://doi.org/10.1007/s10826-015-0351-z

Lobitz, G. K., & Johnson, S. M. (1976). Normal versus deviant children: A multimethod comparison. *Journal of Abnormal Child Psychology*, *3*(4), 353–374. https://doi.org/10.1007/BF00917421

Loe, I. M., & Feldman, H. M. (2007). Academic and Educational Outcomes of Children With ADHD. *Journal of Pediatric Psychology*, *32*(6), 643–654. https://doi.org/10.1093/jpepsy/jsl054

Luman, M., Oosterlaan, J., & Sergeant, J. (2005). The impact of reinforcement contingencies on AD/HD: A review and theoretical appraisal. *Clinical Psychology Review*, *25*(2), 183–213. https://doi.org/10.1016/j.cpr.2004.11.001

Maedgen, J. M., & Carlson, C. L. (2000). Social Functioning and Emotional Regulation in the Attention Deficit Hyperactivity Disorder Subtypes. *Journal of Clinical Child Psychology*, *29*(1), 30–42. https://doi.org/10.1207/S15374424jccp2901_4

March, J. S., Swanson, J. M., Arnold, L. E., Hoza, B., Conners, C. K., Hinshaw, S. P., Hechtman, L., Kraemer, H. C., Greenhill, L. L., Abikoff, H. B., Elliott, L. G., Jensen, P. S., Newcorn, J. H., Vitiello, B., Severe, J., Wells, K. C., & Pelham, W. E. (2000). Anxiety as a predictor and outcome variable in the multimodal treatment study of children with ADHD (MTA). *Journal of Abnormal Child Psychology*, *28*(6), 527–541. https://doi.org/10.1023/a:1005179014321

Mash, E. J., & Hunsley, J. (2005). Evidence-Based Assessment of Child and Adolescent Disorders: Issues and Challenges. *Journal of Clinical Child & Adolescent Psychology*, *34*(3), 362–379. https://doi.org/10.1207/s15374424jccp3403_1

McConaughy, S. H., Volpe, R. J., Antshel, K. M., Gordon, M., & Eiraldi, R. B. (2011). Academic and Social Impairments of Elementary School Children With Attention Deficit Hyperactivity Disorder. *School Psychology Review*, *40*(2), 200–225.

Meinzer, M. C., Lewinsohn, P. M., Pettit, J. W., Seeley, J. R., Gau, J. M., Chronis-Tuscano, A., & Waxmonsky, J. G. (2013). Attention-deficit/hyperactivity disorder in adolescence predicts onset of major depressive disorder through early adulthood. *Depression and Anxiety*, *30*(6), 546–553. https://doi.org/10.1002/da.22082

Mellis, A. M., Woodford, A. E., Stein, J. S., & Bickel, W. K. (2017). A second type of magnitude effect: Reinforcer magnitude differentiates delay discounting between substance users

and controls. *Journal of the Experimental Analysis of Behavior*, *107*(1), 151–160. https://doi.org/10.1002/jeab.235

Merrell, K. W., & Boelter, E. (2001). An Investigation of Relationships Between Social Behavior and ADHD in Children and Youth: Construct Validity of the Home and Community Social Behavior Scales. *Journal of Emotional and Behavioral Disorders*, *9*(4), 260–269. https://doi.org/10.1177/106342660100900406

Mick, E., Byrne, D., Fried, R., Monuteaux, M., Faraone, S. V., & Biederman, J. (2011). Predictors of ADHD persistence in girls at 5-year follow-up. *Journal of Attention Disorders*, *15*(3), 183–192. https://doi.org/10.1177/1087054710362217

Mikami, A. Y., Hinshaw, S. P., Patterson, K. A., & Lee, J. C. (2008). Eating pathology among adolescent girls with attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology*, *117*(1), 225–235. https://doi.org/10.1037/0021-843X.117.1.225

Mowlem, F., Agnew-Blais, J., Taylor, E., & Asherson, P. (2019). Do different factors influence whether girls versus boys meet ADHD diagnostic criteria? Sex differences among children with high ADHD symptoms. *Psychiatry Research*, *272*, 765–773. https://doi.org/10.1016/j.psychres.2018.12.128

Murrell, A. R., Steinberg, D. S., Connally, M. L., Hulsey, T., & Hogan, E. (2015). Acting Out to ACTing On: A Preliminary Investigation in Youth with ADHD and Co-morbid Disorders. *Journal of Child and Family Studies*, *24*(7), 2174–2181. https://doi.org/10.1007/s10826-014-0020-7

Nigg, J. T. (2013). Attention-deficit/hyperactivity disorder and adverse health outcomes. *Clinical Psychology Review*, *33*(2), 215–228. https://doi.org/10.1016/j.cpr.2012.11.005

O'Connor, B. C., Fabiano, G. A., Waschbusch, D. A., Belin, P. J., Gnagy, E. M., Pelham, W. E., Greiner, A. R., & Roemmich, J. N. (2014). Effects of a Summer Treatment Program on Functional Sports Outcomes in Young Children with ADHD. *Journal of Abnormal Child Psychology*, *42*(6), 1005–1017. https://doi.org/10.1007/s10802-013-9830-0

Owens, E. B., & Hinshaw, S. P. (2016). Childhood conduct problems and young adult outcomes among women with childhood attention-deficit/hyperactivity disorder (ADHD). *Journal of Abnormal Psychology*, *125*(2), 220–232. https://doi.org/10.1037/abn0000084

Owens, E. B., Hinshaw, S. P., Kraemer, H. C., Arnold, L. E., Abikoff, H. B., Cantwell, D. P., Conners, C. K., Elliott, G., Greenhill, L. L., Hechtman, L., Hoza, B., Jensen, P. S., March, J. S., Newcorn, J. H., Pelham, W. E., Severe, J. B., Swanson, J. M., Vitiello, B., Wells, K. C., & Wigal, T. (2003). Which treatment for whom for ADHD? Moderators of treatment response in the MTA. *Journal of Consulting and Clinical Psychology*, *71*(3), 540–552. https://doi.org/10.1037/0022-006x.71.3.540

Pastor, P. N., & Reuben, C. A. (2005). Racial and Ethnic Differences in ADHD and LD in Young School-Age Children: Parental Reports in the National Health Interview Survey. *Public Health Reports*, *120*(4), 383–392. https://doi.org/10.1177/003335490512000405

Pastor, P., Reuben, C., Duran, C., & Hawkins, L. (2015). Association between diagnosed ADHD and selected characteristics among children aged 4–17 years: United States, 2011–2013. *National Center for Health Statistics*, *NCHS Data Brief, No 201*.

Pelham, W. E., Burrows-MacLean, L., Gnagy, E. M., Fabiano, G. A., Coles, E. K., Wymbs, B. T., Chacko, A., Walker, K. S., Wymbs, F., Garefino, A., Hoffman, M. T., Waxmonsky, J. G., & Waschbusch, D. A. (2014). A dose-ranging study of behavioral and pharmacological treatment in social settings for children with ADHD. *Journal of Abnormal Child Psychology*, *42*(6), 1019–1031. https://doi.org/10.1007/s10802-013-9843-8

Pelham, W. E., Carlson, C., Sams, S. E., Vallano, G., Dixon, M. J., & Hoza, B. (1993). Separate and combined effects of methylphenidate and behavior modification on boys with attention deficit-hyperactivity disorder in the classroom. *Journal of Consulting and Clinical Psychology*, *61*(3), 506–515. https://doi.org/10.1037/0022-006X.61.3.506

Pelham, W. E., & Fabiano, G. A. (2001). Treatment of attention-deficit hyperactivity disorder: The impact of comorbidity. *Clinical Psychology & Psychotherapy*, *8*(5), 315–329. https://doi.org/10.1002/cpp.312

Pelham, W. E., & Fabiano, G. A. (2008). Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *37*(1), 184–214. https://doi.org/10.1080/15374410701818681

Pelham, 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 & Adolescent Psychology*, *34*(3), 449–476. https://doi.org/10.1207/s15374424jccp3403_5

Pelham, W. E., Fabiano, G. A., Waxmonsky, J. G., Greiner, A. R., Gnagy, E. M., Pelham, W. E., Coxe, S., Verley, J., Bhatia, I., Hart, K., Karch, K., Konijnendijk, E., Tresco, K., Nahum-Shani, I., & Murphy, S. A. (2016). Treatment Sequencing for Childhood ADHD: A Multiple-Randomization Study of Adaptive Medication and Behavioral Interventions. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *45*(4), 396–415. https://doi.org/10.1080/15374416.2015.1105138

Pelham, W. E., Gnagy, E. M., Grenier, 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.

Pelham, W. E., Greiner, A., R., & Gnagy, E. M. (2017). *Children's Summer Treatment Program Manual*.

Pelham, W. E., & Hoza, B. (1996). Intensive treatment: A summer treatment program for children with ADHD. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice.* (pp. 311–340). American Psychological Association. https://doi.org/10.1037/10196-013

Piper, B. J., Gray, H. M., Raber, J., & Birkett, M. A. (2014). Reliability and validity of Brief Problem Monitor, an abbreviated form of the Child Behavior Checklist. *Psychiatry and Clinical Neurosciences*, *68*(10), 759–767. https://doi.org/10.1111/pcn.12188

Pituch, K., & Stevens, J. (2015). Repeated-Measures Analysis. In *Applied Multivariate Statistics for the Social Sciences: Analyses with SAS and IBM's SPSS, Sixth Edition* (6th ed., pp. 471-536). Taylor & Francis Group.

Power, T. J., Watkins, M. W., Anastopoulos, A. D., Reid, R., Lambert, M. C., & DuPaul, G. J. (2017). Multi-Informant Assessment of ADHD Symptom-Related Impairments Among Children and Adolescents. *Journal of Clinical Child & Adolescent Psychology*, *46*(5), 661-674. https://doi.org/10.1080/15374416.2015.1079781

Poznanski, B., Hart, K. C., & Cramer, E. (2018). Are Teachers Ready? Preservice Teacher Knowledge of Classroom Management and ADHD. *School Mental Health*, *10*(3), 301-313. https://doi.org/10.1007/s12310-018-9259-2

Ramtekkar, U. P., Reiersen, A. M., Todorov, A. A., & Todd, R. D. (2010). Sex and age differences in attention-deficit/hyperactivity disorder symptoms and diagnoses: Implications for DSM-V and ICD-11. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(3), 217-228.e1-3.

Rapport, M. D., Murphy, H. A., & Bailey, J. S. (1982). Ritalin vs. response cost in the control of hyperactive children: A within-subject comparison. *Journal of Applied Behavior Analysis*, *15*(2), 205-216. https://doi.org/10.1901/jaba.1982.15-205

Rapport, M. D., Scanlan, S. W., & Denney, C. B. (1999). Attention-Deficit/Hyperactivity Disorder and Scholastic Achievement: A Model of Dual Developmental Pathways. *Journal of Child Psychology and Psychiatry*, *40*(8), 1169-1183. https://doi.org/10.1111/1469-7610.00534

Riglin, L., Leppert, B., Dardani, C., Thapar, A. K., Rice, F., O'Donovan, M. C., Davey Smith, G., Stergiakouli, E., Tilling, K., & Thapar, A. (2020). ADHD and depression: Investigating a causal explanation. *Psychological Medicine*, 1–8. https://doi.org/10.1017/S0033291720000665

Rowland, A. S., Skipper, B. J., Umbach, D. M., Rabiner, D. L., Campbell, R. A., Naftel, A. J., & Sandler, D. P. (2015). The Prevalence of ADHD in a Population-Based Sample. *Journal of Attention Disorders*, *19*(9), 741–754. https://doi.org/10.1177/1087054713513799

Rowland, A. S., Umbach, D. M., Stallone, L., Naftel, A. J., Bohlig, E. M., & Sandler, D. P. (2002). Prevalence of medication treatment for attention deficit-hyperactivity disorder among elementary school children in Johnston County, North Carolina. *American Journal of Public Health*, *92*(2), 231–234. https://doi.org/10.2105/ajph.92.2.231

Roy, A., Hechtman, L., Arnold, L. E., Sibley, M. H., Molina, B. S. G., Swanson, J. M., Howard, A. L., Vitiello, B., Severe, J. B., Jensen, P. S., Arnold, L. E., Hoagwood, K., Richters, J., Vereen, D., Hinshaw, S. P., Elliott, G. R., Wells, K. C., Epstein, J. N., Murray, D. W., … Stern, K. (2016). Childhood Factors Affecting Persistence and Desistence of Attention-Deficit/Hyperactivity Disorder Symptoms in Adulthood: Results From the MTA. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(11), 937–944.e4. https://doi.org/10.1016/j.jaac.2016.05.027

Rzepa, S. R., & Marocco, M. (2008, February). *Development, Reliability, and Validity of the Conners 3 Short Forms.* Annual Meeting of the National Association of School Psychologists, New Orleans, LA.

Sagvolden, T., Aase, H., Zeiner, P., & Berger, D. (1998). Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behavioural Brain Research*, *94*(1), 61–71. https://doi.org/10.1016/S0166-4328(97)00170-8

Sasser, T. R., Kalvin, C. B., & Bierman, K. L. (2016). Developmental trajectories of clinically significant attention-deficit/hyperactivity disorder (ADHD) symptoms from grade 3 through 12 in a high-risk sample: Predictors and outcomes. *Journal of Abnormal Psychology*, *125*(2), 207–219. https://doi.org/10.1037/abn0000112

Scaramella, L. V., & Leve, L. D. (2004). Clarifying Parent–Child Reciprocities During Early Childhood: The Early Childhood Coercion Model. *Clinical Child and Family Psychology Review*, *7*(2), 89–107. https://doi.org/10.1023/B:CCFP.0000030287.13160.a3

Scheffler, R. M., Brown, T. T., Fulton, B. D., Hinshaw, S. P., Levine, P., & Stone, S. (2009). Positive Association Between Attention-Deficit/ Hyperactivity Disorder Medication Use and Academic Achievement During Elementary School. *PEDIATRICS*, *123*(5), 1273–1279. https://doi.org/10.1542/peds.2008-1597

Sibley, M. H., Coxe, S. J., Campez, M., Morley, C., Olson, S., Hidalgo-Gato, N., Gnagy, E., Greiner, A., Coles, E. K., Page, T., & Pelham, W. E. (2018). High versus Low Intensity Summer Treatment for ADHD Delivered at Secondary School Transitions. *Journal of Clinical Child & Adolescent Psychology*, *47*(2), 248–265. https://doi.org/10.1080/15374416.2018.1426005

Sibley, M. H., Swanson, J. M., Arnold, L. E., Hechtman, L. T., Owens, E. B., Stehli, A., Abikoff, H., Hinshaw, S. P., Molina, B. S. G., Mitchell, J. T., Jensen, P. S., Howard, A. L., Lakes, K. D.,

Pelham, W. E., & MTA Cooperative Group. (2017). Defining ADHD symptom persistence in adulthood: Optimizing sensitivity and specificity. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *58*(6), 655–662. https://doi.org/10.1111/jcpp.12620

Siebelink, N. M., Bögels, S. M., Speckens, A. E. M., Dammers, J. T., Wolfers, T., Buitelaar, J. K., & Greven, C. U. (2022). A randomised controlled trial (MindChamp) of a mindfulness-based intervention for children with ADHD and their parents. *Journal of Child Psychology and Psychiatry*, *63*(2), 165–177. https://doi.org/10.1111/jcpp.13430

Skinner, B. F. (1963). Operant behavior. *American Psychologist*, *18*(8), 503–515. https://doi.org/10.1037/h0045185

Sleator, E. K., & Ullmann, R. K. (1981). Can the physician diagnose hyperactivity in the office? *Pediatrics*, *67*(1), 13–17.

Sonuga-Barke, E. J. S. (2002). Psychological heterogeneity in AD/HD—a dual pathway model of behaviour and cognition. *Behavioural Brain Research*, *130*(1–2), 29–36. https://doi.org/10.1016/S0166-4328(01)00432-6

Sonuga-Barke, E. J. S. (2003). The dual pathway model of AD/HD: An elaboration of neuro-developmental characteristics. *Neuroscience & Biobehavioral Reviews*, *27*(7), 593–604. https://doi.org/10.1016/j.neubiorev.2003.08.005

Sonuga-Barke, E. J. S., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., Stevenson, J., Danckaerts, M., van der Oord, S., Döpfner, M., Dittmann, R. W., Simonoff, E., Zuddas, A., Banaschewski, T., Buitelaar, J., Coghill, D., Hollis, C., Konofal, E., Lecendreux, M., … European ADHD Guidelines Group. (2013). Nonpharmacological Interventions for ADHD:

Systematic Review and Meta-Analyses of Randomized Controlled Trials of Dietary and Psychological Treatments. *American Journal of Psychiatry*, *170*(3), 275–289. https://doi.org/10.1176/appi.ajp.2012.12070991

Sonuga-Barke, E. J. S., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-Based Therapies for Preschool Attention-Deficit/Hyperactivity Disorder: A Randomized, Controlled Trial With a Community Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(4), 402–408. https://doi.org/10.1097/00004583-200104000-00008

Southam-Gerow, M. A., & Prinstein, M. J. (2014). Evidence Base Updates: The Evolution of the Evaluation of Psychological Treatments for Children and Adolescents. *Journal of Clinical Child & Adolescent Psychology*, *43*(1), 1–6. https://doi.org/10.1080/15374416.2013.855128

Spencer, T., Biederman, J., Wilens, T., Harding, M., O'Donnell, D., & Griffin, S. (1996). Pharmacotherapy of Attention-Deficit Hyperactivity Disorder across the Life Cycle. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(4), 409–432. https://doi.org/10.1097/00004583-199604000-00008

Staddon, J. E. R., & Cerutti, D. T. (2003). Operant Conditioning. *Annual Review of Psychology*, *54*(1), 115–144. https://doi.org/10.1146/annurev.psych.54.101601.145124

Steele, M., Jensen, P. S., & Quinn, D. M. P. (2006). Remission versus response as the goal of therapy in ADHD: A new standard for the field? *Clinical Therapeutics*, *28*(11), 1892–1908. https://doi.org/10.1016/j.clinthera.2006.11.006

Swanson, J. M., Gupta, S., Williams, L., Agler, D., Lerner, M., & Wigal, S. (2002). Efficacy of a New Pattern of Delivery of Methylphenidate for the Treatment of ADHD: Effects on Activity Level in the Classroom and on the Playground. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(11), 1306–1314. https://doi.org/10.1097/00004583-200211000-00011

Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., Clevenger, W., Davies, M., Elliott, G. R., Greenhill, L. L., Hechtman, L., Hoza, B., Jensen, P. S., March, J. S., Newcorn, J. H., Owens, E. B., Pelham, W. E., Schiller, E., Severe, J. B., … Wu, M. (2001). Clinical Relevance of the Primary Findings of the MTA: Success Rates Based on Severity of ADHD and ODD Symptoms at the End of Treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(2), 168–179. https://doi.org/10.1097/00004583-200102000-00011

Synn, S., Tannenbaum, S., Beraki, Y., & Wilson, M. (2019). *Camp Baker Staff Program Manual: An Adapted Summer Treatment Program*.

Tannenbaum, S., Kim, R., Synn, S., Beraki, Y., & Wilson, M. (2019). *A Pilot Study of Camp Baker, an Adapted Summer Treatment Program for Children with ADHD.*

Thomas, R., Sanders, S., Doust, J., Beller, E., & Glasziou, P. (2015). Prevalence of Attention-Deficit/Hyperactivity Disorder: A Systematic Review and Meta-analysis. *PEDIATRICS*, *135*(4), e994–e1001. https://doi.org/10.1542/peds.2014-3482

Thorell, L. B. (2007). Do delay aversion and executive function deficits make distinct contributions to the functional impact of ADHD symptoms? A study of early academic

skill deficits. *Journal of Child Psychology and Psychiatry*, *48*(11), 1061–1070. https://doi.org/10.1111/j.1469-7610.2007.01777.x

Tung, I., Li, J. J., Meza, J. I., Jezior, K. L., Kianmahd, J. S. V., Hentschel, P. G., ONeil, P. M., & Lee, S. S. (2016). Patterns of Comorbidity Among Girls With ADHD: A Meta-analysis. *PEDIATRICS*, *138*(4), e20160430–e20160430. https://doi.org/10.1542/peds.2016-0430

Turgay, A., Goodman, D. W., Asherson, P., Lasser, R. A., Babcock, T. F., Pucci, M. L., Barkley, R., & ADHD Transition Phase Model Working Group. (2012). Lifespan persistence of ADHD: The life transition model and its application. *The Journal of Clinical Psychiatry*, *73*(2), 192–201. https://doi.org/10.4088/JCP.10m06628

Unnever, J. D., & Cornell, D. G. (2003). Bullying, Self-Control, and Adhd. *Journal of Interpersonal Violence*, *18*(2), 129–147. https://doi.org/10.1177/0886260502238731

van den Hoofdakker, B. J., Hoekstra, P. J., van der Veen-Mulders, L., Sytema, S., Emmelkamp, P. M. G., Minderaa, R. B., & Nauta, M. H. (2014). Paternal influences on treatment outcome of behavioral parent training in children with attention-deficit/hyperactivity disorder. *European Child & Adolescent Psychiatry*, *23*(11), 1071–1079. https://doi.org/10.1007/s00787-014-0557-4

van den Hoofdakker, B. J., Nauta, M. H., Veen-Mulders, L. v. d., Sytema, S., Emmelkamp, P. M. G., Minderaa, R. B., & Hoekstra, P. J. (2010). Behavioral Parent Training as an Adjunct to Routine Care in Children with Attention-Deficit/Hyperactivity Disorder: Moderators of

Treatment Response. *Journal of Pediatric Psychology*, *35*(3), 317–326. https://doi.org/10.1093/jpepsy/jsp060

Van der Oord, S., Prins, P. J. M., Oosterlaan, J., & Emmelkamp, P. M. G. (2008). Efficacy of methylphenidate, psychosocial treatments and their combination in school-aged children with ADHD: A meta-analysis. *Clinical Psychology Review*, *28*(5), 783–800. https://doi.org/10.1016/j.cpr.2007.10.007

Visser, S., Danielson, M., Wolraich, M., Fox, M., Frosse, S., Valle, L., Holbrook, J., Claussen, A., & Peacock, G. (2016). Vital signs: National and state-specific patterns of attention deficit/hyperactivity disorder treatment among insured children aged 2–5 years—United States, 2008–2014. *Morbidity and Mortality Weekly Report, US Department of Health and Human Services / Centers for Disease Control and Prevention*, *65*(17), 443–450.

Volpe, R. J., & Fabiano, G. A. (2013). *Daily behavior report cards: An evidence-based system of assessment and intervention*. The Guilford Press.

Wells, K. C., Pelham, W. E., Kotkin, R. A., Hoza, B., Abikoff, H. B., Abramowitz, A., Arnold, L. E., Cantwell, D. P., Conners, C. K., Del Carmen, R., Elliott, G., Greenhill, L. L., Hechtman, L., Hibbs, E., Hinshaw, S. P., Jensen, P. S., March, J. S., Swanson, J. M., & Schiller, E. (2000). Psychosocial treatment strategies in the MTA study: Rationale, methods, and critical issues in design and implementation. *Journal of Abnormal Child Psychology*, *28*(6), 483–505. https://doi.org/10.1023/a:1005174913412

Wilens, T. E. (2008). Effects of Methylphenidate on the Catecholaminergic System in Attention-Deficit/Hyperactivity Disorder: *Journal of Clinical Psychopharmacology*, *28*(Suppl. 2), S46–S53. https://doi.org/10.1097/JCP.0b013e318173312f

Willcutt, E. G. (2012). The Prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. *Neurotherapeutics*, *9*(3), 490–499. https://doi.org/10.1007/s13311-012-0135-8

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the Executive Function Theory of Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. *Biological Psychiatry*, *57*(11), 1336–1346. https://doi.org/10.1016/j.biopsych.2005.02.006

Williford, A. P., & Vitiello, V. E. (2020). Who's in charge? Child behavior predicts teacher subsequent classroom management practice for preschoolers reported to display disruptive behavior. *School Psychology*, *35*(5), 299–310. https://doi.org/10.1037/spq0000402

Wolraich, M. L., Hagan, J. F., Allan, C., Chan, E., Davison, D., Earls, M., Evans, S. W., Flinn, S. K., Froehlich, T., Frost, J., Holbrook, J. R., Lehmann, C. U., Lessin, H. R., Okechukwu, K., Pierce, K. L., Winner, J. D., Zurhellen, W., & SUBCOMMITTEE ON CHILDREN AND ADOLESCENTS WITH ATTENTION-DEFICIT/HYPERACTIVE DISORDER. (2019). Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents. *Pediatrics*, *144*(4), e20192528. https://doi.org/10.1542/peds.2019-2528

Wolraich, M. L., McKeown, R. E., Visser, S. N., Bard, D., Cuffe, S., Neas, B., Geryk, L. L., Doffing, M., Bottai, M., Abramowitz, A. J., Beck, L., Holbrook, J. R., & Danielson, M. (2014). The Prevalence of ADHD: Its Diagnosis and Treatment in Four School Districts Across Two States. *Journal of Attention Disorders*, *18*(7), 563–575. https://doi.org/10.1177/1087054712453169

Wright, N., Moldavsky, M., Schneider, J., Chakrabarti, I., Coates, J., Daley, D., Kochhar, P., Mills, J., Sorour, W., & Sayal, K. (2015). Practitioner Review: Pathways to care for ADHD - a systematic review of barriers and facilitators. *Journal of Child Psychology and Psychiatry*, *56*(6), 598–617. https://doi.org/10.1111/jcpp.12398

Yamashita, Y., Mukasa, A., Anai, C., Honda, Y., Kunisaki, C., Koutaki, J., Tada, Y., Egami, C., Kodama, N., Nakashima, M., Nagamitsu, S., & Matsuishi, T. (2011). Summer treatment program for children with attention deficit hyperactivity disorder: Japanese experience in 5 years. *Brain and Development*, *33*(3), 260–267. https://doi.org/10.1016/j.braindev.2010.09.005

Yamashita, Y., Mukasa, A., Honda, Y., Anai, C., Kunisaki, C., Koutaki, J., Motoyama, S., Miura, N., Sugimoto, A., Ohya, T., Nakashima, M., Nagamitsu, S., Gnagy, E. M., Greiner, A. R., Pelham, W. E., & Matsuishi, T. (2010). Short-term effect of American summer treatment program for Japanese children with attention deficit hyperactivity disorder. *Brain and Development*, *32*(2), 115–122. https://doi.org/10.1016/j.braindev.2008.12.004

Yeh, M., McCabe, K., Hurlburt, M., Hough, R., Hazen, A., Culver, S., Garland, A., & Landsverk, J. (2002). Referral sources, diagnoses, and service types of youth in public outpatient

mental health care: A focus on ethnic minorities. *The Journal of Behavioral Health Services & Research*, *29*(1), 45–60. https://doi.org/10.1007/BF02287831

Young, S., & Gudjonsson, G. H. (2008). Growing out of ADHD: The relationship between functioning and symptoms. *Journal of Attention Disorders*, *12*(2), 162–169. https://doi.org/10.1177/1087054707299598

Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M.-L., & Youngstrom, J. K. (2017). Evidence-Based Assessment as an Integrative Model for Applying Psychological Science to Guide the Voyage of Treatment. *Clinical Psychology: Science and Practice*, *24*(4), 331–363. https://doi.org/10.1111/cpsp.12207