PARENT-OTHER AND PARENT-CHILD RATINGS AGREEMENT ON ADHD AND CO-OCCURING EMOTIONAL AND BEHAVIOR SYMPTOMS AMONG CHILDREN WITH ADHD IN THE MADDY STUDY

ANUPAM ROY
Master of Science, University of Regina, 2017

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

HEALTH SCIENCES (PUBLIC HEALTH)

Faculty of Health Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

PARENT-OTHER AND PARENT-CHILD RATINGS AGREEMENT ON ADHD AND CO-OCCURING EMOTIONAL AND BEHAVIOR SYMPTOMS AMONG CHILDREN WITH ADHD IN THE MADDY STUDY

ANUPAM ROY

Date of Defense: May 26, 2026

Dr. Brenda Leung Thesis Supervisor	Associate Professor	Ph.D.
Dr. Cindy Feng Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. James Sanders Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. Julia Brassolotto Chair, Thesis Examination Committee	Associate Professor	Ph.D.

DEDICATION

I dedicate this thesis to my wife, whose unwavering love, patience, and encouragement sustained me throughout this journey. Your constant support, understanding, and belief in me gave me the strength to persevere through every challenge. This achievement is as much yours as it is mine.

ABSTRACT

Background: Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental condition commonly diagnosed in childhood. Multi-informant (parents, teachers, and children) reporting of symptoms is a standard practice in pediatric behavior research. However, informant discrepancies are frequently observed and may influence screening accuracy. The objective of this study was to examine the level of parent-other adult (teacher, grandparent, aunt/uncle, or neighbor who knew the child's behavior well) and parent-child agreement for ADHD core symptoms and co-occurring emotional and behavior symptoms in children with ADHD and whether the agreement differs by child sex.

Methods: Secondary analysis of data from the MADDY study was conducted to assess the level of parent-other adult and parent-child ratings agreement. Parent-other adult agreement was assessed across CASI-5 domains, including inattention, hyperactivity, social and academic impairment, oppositional defiant disorder (ODD), disruptive mood dysregulation disorder (DMDD), and anger related behavior. Parent-child agreement was assessed across PROMIS domains, including anger, anxiety, depression, peer relations, and sleep. Agreement at all domains were analyzed using data of two timepoints: baseline and week 8. Descriptive statistics summarized demographic and clinical characteristics of study participants. Informant agreement was assessed using concordance correlation coefficients (CCCs), with confidence intervals estimated via bootstrapping. Bland-Altman plots were used to visually assess systematic bias, identify patterns of disagreement, and evaluate the limits of agreement across raters. Stratified analyses by child sex were conducted to explore sex-based differences in informant agreement.

Results: A total of 126 children were enrolled in the RCT (mean age = 9.8 years; 73% boys), with analytic sample sizes for this study varying by domain due to data availability. Most

families (57.6%) reported annual household incomes > $80,000, and 61.1% parents held a university degree or higher. Parent-other adult agreement on CASI-5 domains was poor at baseline (CCC = 0.12 - 0.43) and week 8 (CCC = 0.34 - 0.53), with agreement for inattention significantly improved at week 8 (diff CCC = -0.27 (95% CI: -0.44 to -0.09)). Parent-child agreement across PROMIS domains was also poor at baseline (CCC = 0.07 - 0.39) and week 8 (CCC = 0.09 - 0.26). Parents tended to report higher symptom severity than other adults across most CASI-5 domains at both time points, and greater anger severity than their children at both time points as well as lower sleep disturbance at week 8. Sex-specific differences in agreement were observed, with higher parent-other adult agreement for inattention among girls at week 8 (diff CCC = -0.33 (95% CI: -0.58 to -0.04)) and higher parent-child agreement for peer relations among boys at baseline (diff CCC = 0.46 (95% CI: 0.08 to 0.81)).

Conclusions: Overall low level of parent-other adult and parent-child ratings agreement underscores the differences in informant perspectives. It also reflects variations in contextual demands, informants' knowledge and expectations, the nature of their relationships with the child, and subjective interpretations of behavioral and emotional symptoms. Children's sex may play a role in symptom recognition and reporting patterns, and further studies stratified by sex is warranted. Overall, these findings highlight the importance of incorporating sex-sensitive framework and multi-informant assessment approaches to improve the precision of screening in research context.

Keywords: ADHD, child's emotional and behavioral problems, child sex differences, informant agreement.

USE OF GENERATIVE AI

I would like to acknowledge the use of AI-powered tools, including Microsoft Word Editor and ChatGPT, to support the grammar and writing style of this work. These tools have been useful in detecting grammatical error, improving clarity and readability, particularly when simplifying or condensing lengthy sentences. However, all ideas, analyses, and interpretations presented in this work are entirely my own.

ACKNOWLEDGEMENTS

With sincere gratitude, I acknowledge the individuals and institutions who supported me throughout this challenging and rewarding journey.

First and foremost, I am deeply grateful to my supervisor, Dr. Brenda Leung, whose unwavering support and mentorship made this research possible. Thank you for your invaluable guidance, thoughtful feedback, and continuous encouragement. Your expertise and insight were instrumental in shaping this project and greatly enriched my academic experience. I am sincerely thankful for the opportunity to collaborate with you and will carry the lessons I have gained into my future endeavors.

My sincere thanks also go to my supervisory committee member, Dr. Cindy Feng, for her guidance and expertise throughout the analytical stages of this project. Your constructive feedback and insightful suggestions significantly strengthened the rigor and quality of this thesis. I am truly grateful for the time and support you provided throughout this process.

My gratitude also extends to supervisory committee member, Dr. James Sanders, for his expertise and valuable input. Your perspectives and feedback have meaningfully contributed to the development and improvement of this project.

I am also thankful to the MADDY team for providing the opportunity to use the data for this study. Additionally, I gratefully acknowledge the financial support provided by the University of Lethbridge. Receiving this support was an honor, and it played a crucial role in enabling the completion of this research.

Finally, I am deeply thankful to my family for their unconditional love, encouragement, and sacrifices. Your constant support has been the backbone of all my accomplishments, and reaching this milestone would not have been possible without you.

TABLE OF CONTENTS

DEDICATION iii
ABSTRACT iv
USE OF GENERATIVE AI vi
ACKNOWLEDGEMENTS vii
LIST OF TABLES xi
LIST OF FIGURES xii
LIST OF ABBREVIATIONS xiii
CHAPTER 1: INTRODUCTION 1
1.1 Attention-Deficit/Hyperactivity Disorder 1
1.2 Risk Factors and Impact of ADHD in Children 2
1.3 Medication Risks & Nutrient Benefits 4
1.4 ADHD Diagnosis and Screening in Pediatric Research 6
1.5 Informant Discrepancies and Multi-Informant Perspectives 7
1.6 Parent-Teacher Agreement in ADHD Symptoms and Related Disorders 8
1.7 Parent-Child Agreement in Internalizing and Externalizing Symptoms 10
1.8 Difference in Informant Agreement by Child Gender or Sex 12
1.9 Study Rationale 13
1.10 Research Objective 14
1.11 Research Questions 14
CHAPTER 2: METHODS 16
2.1 Study Design 16
2.2 Setting 16
2.3 Participants 17
2.3.1 Eligibility Criteria 17
2.3.2 Source and Selection Method 18
2.4 Data Collection 19
2.5 Sample Size 20
2.6 Measures 20
2.6.1 Child & Adolescent Symptom Inventory–5 (CASI-5) 20
2.6.2 Patient-Reported Outcomes Measurement Information System (PROMIS) 21
2.6.3 Demographic Characteristics 23
2.6.4 Clinical Characteristics 24
CHAPTER 3: ANALYSIS STRATEGY 25
3.1 Sample Description 25

3.2 Agreement Assessment 25
3.3 Missing Data 28
CHAPTER 4: RESULT 29
4.1 Participant Characteristics 29
4.2 Concordance Correlation Coefficients (CCCs) for Parent–Other Adult Reports on CASI-5 Domains 29
4.3 CCCs for Parent–Child Reports on PROMIS Domains 29
4.4 Sex-Stratified Analyses 30
4.5 Summary of Informant Ratings 31
4.6 Summary of Informant Ratings by Child Sex 32
4.7 Bland-Altman Plots 33
4.7.1 Parent-Other Adult Agreement (CASI-5) 33
4.7.2 Parent-Child Agreement (PROMIS) 36
CHAPTER 5: DISCUSSION 48
5.1 Parent-Other Adult Rating Agreement 48
5.1.1 Inattention and Hyperactivity 49
5.1.2 Impairment 53
5.1.3 ODD 55
5.1.4 DMDD 57
5.1.5 Anger 58
5.2 Parent-Child Rating Agreement 59
5.2.1 Anger 59
5.2.2 Anxiety 61
5.2.3 Depression 62
5.2.4 Peer Relations 63
5.2.5 Sleep 65
5.3 Agreement By Child Sex 66
5.3.1 CASI-5 Domains 66
5.3.2 PROMIS Domains 68
5.4 Strength and Limitations 70
5.5 Conclusion 72
REFERENCES 74

LIST OF TABLES

Chapter 4

Table 1: Socio demographics and clinical characteristics of study participants ……..…….… 38

Table 2: CASI-5 Parent-Other Adult Agreement …… 41

Table 3: PROMIS Parent-Child Agreement …… 41

Table 4: Baseline and Week 8 CASI-5 by child sex …… 42

Table 5: Baseline and Week 8 PROMIS by child sex ……...…….……. 43

Table 6: Summary of parent-other adult ratings CASI-5 Domains ……...……...….... 44

Table 7: Summary of parent-child ratings PROMIS Domains …… 44

Table 8: Summary of parent-other adult ratings CASI-5 domains by child sex …….. 45

Table 9: Summary of parent-child ratings PROMIS domains by child sex ……..... 45

LIST OF FIGURES

Chapter 4

Figure1: Bland-Altman Plots for Parent-Other Adult Agreement on CASI-5 Domains ……. 46

Figure 2: Bland-Altman Plots for Parent-Child Agreement on PROMIS Domains ……..…..... 47

LIST OF ABBREVIATIONS

ADHD: Attention-Deficit/Hyperactivity Disorder

CASI-5: Child & Adolescent Symptom Inventory-5

CCC: Concordance Correlation Coefficients

DMDD: Disruptive Mood Dysregulation Disorder

MADDY: Micronutrient for Attention-Deficit/Hyperactivity Disorder in Youth

ODD: Oppositional defiant disorder

PROMIS: Patient-Reported Outcomes Measurement Information System

CHAPTER 1: INTRODUCTION

1.1 Attention-Deficit/Hyperactivity Disorder

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder affecting approximately 8% of children worldwide (Salari et al., 2023; Ayano et al., 2023). This disorder is characterised by a persistent pattern of impaired attention, impulsivity, and hyperactivity. Three subtypes of ADHD are classified as inattentive ADHD, hyperactive/impulsive ADHD, and a combined type (American Psychiatric Association, 2022). Among these, the inattentive type is the most prevalent, followed by the hyperactive/impulsive and the combined type (Ayano et al., 2023). According to World Health Organization, inattention refers to having trouble staying focused on tasks that do not provide high level of stimulation or frequent rewards, getting easily distracted, and struggling with organization. Hyperactivity involves being overly active and having difficulty sitting still, especially in situations that require self-control. Impulsivity is a tendency of acting quickly without thinking about the risks or consequences (World Health Organization [WHO], 2019).

ADHD is generally diagnosed in childhood and persists into adulthood. Its impact varies by age and gender. Hyperactivity is usually hard to distinguish from developmental factors before age four, but the condition tends to be evident in elementary school (Salari et al., 2023). Studies also showed that the main characteristics of ADHD change as people grow older, and regardless of classification the symptom severity often decreases in adulthood (Salari et al., 2023). Boys are diagnosed with ADHD at nearly twice the rate as girls, with prevalence estimate of 10% in boys and 5% in girls (Ayano et al., 2023). According to a study by Rucklidge (2010), the condition presents differently in boys and girls. Boys with ADHD often exhibit externalized behaviors,

such as hyperactivity and impulsivity, while girls tend to display internalized symptoms, including inattentiveness and low self-esteem. Additionally, Rucklidge (2010) emphasized that boys are generally more physically aggressive and outwardly express frustration, whereas girls are more verbally aggressive and internalize their emotions, increasing their risk of developing anxiety, depression, and eating disorders. ADHD in girls may be underdiagnosed partly because they are more likely to exhibit the inattentive subtype rather than other subtypes (Ayano et al., 2023; Quinn & Madhoo, 2014), making their behavior less noticeable and less often seen as disruptive (Gilbert et al., 2025). Together, these patterns underscore the importance of understanding both the risk factors associated with ADHD and its wide-ranging impacts on children's development.

1.2 Risk Factors and Impact of ADHD in Children

ADHD in children is influenced by various factors, including genetic and environmental components (Azeredo et al., 2018; Kim et al., 2020). Within the family environment, the quality of parent-child interactions, characterized by parental sensitivity, levels of intrusiveness, and the use of harsh or inconsistent discipline, has been associated with ADHD outcomes. Additionally, adverse family circumstances, such as maltreatment or physical abuse, parental divorce or single parenting, parental incarceration, and high levels of child media exposure are also associated with an increased risk of ADHD in children (Claussen et al., 2024). Biological factors in the prenatal period also contribute significantly to ADHD risk. Maternal health conditions, including obesity, overweight, preeclampsia, and hypertension, as well as maternal acetaminophen exposure or smoking during pregnancy, and childhood atopic diseases were strongly associated with ADHD in offspring (Kim et al., 2020). These genetic, biological, and environmental risk factors may arise simultaneously. Evidence suggests that children exposed to multiple risk

factors at the same time face higher likelihood of developing ADHD symptoms (Huhdanpää et al., 2021).

ADHD affects multiple domains of a child's functioning. Children with ADHD frequently experience academic difficulties that interfere with classroom learning, task completion, and overall academic achievement (Arnold et al., 2020; DuPaul et al., 2004; McConaughy et al., 2011). In addition, ADHD has a substantial impact on social functioning, with affected children often struggling to form and maintain friendships and being more vulnerable to peer rejection, bullying, and interpersonal conflict (McConaughy et al., 2011; Sciberras et al., 2009). These social difficulties are likely associated with the aggressive or disruptive behaviors observed in children with ADHD (McConaughy et al., 2011). Based on parent and/or teacher ratings, research showed that children without ADHD exhibit significantly better academic performance and social functioning compared to children with ADHD (DuPaul et al., 2004; Faraone et al., 1998; McConaughy et al., 2011).

ADHD in children and adolescents is associated with mental health difficulties, with affected children showing higher rates of oppositional defiant disorder (ODD) and conduct disorder (CD), and higher risk of antisocial personality disorder in adulthood (Sciberras et al., 2009). Furthermore, emerging evidence indicates that ADHD is one of the leading contributors to the global disease burden among children and adolescents. The cost of raising children with ADHD is significantly higher compared to raising children without such disorder (Chan et al., 2002; Doshi et al., 2012). Besides the core symptoms of ADHD, children often experience comorbid conditions such as anxiety, depression, anger, irritability, emotional dysregulation, ODD, CD, peer relationship difficulties, learning disabilities, and sleep disturbances, which has a direct negative impact on academic, occupational, emotional, or social functioning (Dickstein, 2016;

Gnanavel et al., 2019; Lenzi et al., 2018). Previous studies have also shown that early identification and intervention for ADHD and related problem behaviors can reduce the risk of long-term adverse outcomes (Sonuga-Barke et al., 2011). A range of pharmacological and non-pharmacological treatments has been explored for managing ADHD; however, concerns regarding medication-related risks and increasing interest in nutritional approaches highlight the need to better understand alternative or complementary strategies.

1.3 Medication Risks & Nutrient Benefits

Various interventions for ADHD have been studied, including medication, psychosocial interventions, parent support, nutrition and supplements, neurofeedback, neurostimulation, physical exercise, complementary medicine, and school interventions (Peterson et al., 2024b). Among these, pharmacologic interventions consisting of stimulant and non-stimulant medications are most often used to address ADHD symptoms. While these medications can improve symptoms for many individuals, they can cause a range of adverse effects, varying from mild to severe. For example, a meta-analysis by Hennissen et al. (2017) found that medications such as methylphenidate, amphetamines, and atomoxetine led to slight but statistically significant increases in blood pressure and heart rate, which are the risk factors for cardiovascular disease. Additionally, a nationwide cohort study in Denmark highlighted that stimulant treatment in children with ADHD was associated with an increased risk of cardiovascular events such as arrhythmia, cerebrovascular disease, hypertension, ischemic heart disease, heart failure, and pulmonary hypertension, with a complex, time-dependent dose-response relationship (Dalsgaard et al., 2014). A review of longitudinal studies suggests that treating ADHD children with stimulant medication results in a reduction in both height and weight gain (Faraone et al., 2008). Other common side effects of stimulant medications in

ADHD children include trouble sleeping, headache, dizziness, mood swings, irritability, anxiety, dermatological reactions, ocular, or gastrointestinal symptoms, (Bingöl-Kızıltunç et al., 2022; Graham et al., 2011; Graham and Coghill, 2008; Kamal et al., 2018; Nanda et al., 2023). Growing concerns of sides of effects of medications have led to families' interest in using broad-spectrum multinutrient supplementation to manage ADHD and related mood and behavior issues.

Emerging research indicates that multinutrient supplementation offers significant benefits in improving ADHD symptoms, quality of life, mood issues or emotional dysregulation, with small-to-medium effects (Gordon et al., 2015; Rucklidge et al., 2014; Rucklidge et al., 2011; Rucklidge et al., 2018). RCTs conducted in New Zealand (Rucklidge et al., 2014; Rucklidge et al., 2018) demonstrated that multinutrient treatment led to greater overall symptom improvement in both children and adults with ADHD, as measured by the Clinical Global Impression-Improvement (CGI-I) scale, with no serious adverse events reported. In children, clinicians reported a clinically meaningful improvement in inattentive symptoms with multinutrient treatment compared to placebo (OR = 4.9; 95% CI: 1.5–16.3). Additionally, reports from clinicians, parents, and teachers indicated better emotional regulation, reduced aggression, and improved general functioning in the multinutrient group (Rucklidge et al., 2018). In adults with moderate to severe depression, multinutrient treatment was associated with greater mood improvements (Rucklidge et al., 2014). Moreover, significant improvements were observed across self, observer, and clinician reports on inattention, hyperactivity/impulsivity, mood, quality of life, anxiety, and stress (Rucklidge et al., 2011).

The Micronutrient for Attention-Deficit/Hyperactivity Disorder in Youth (MADDY) study, a multi-site, blinded, placebo-controlled RCT in North America also supported these findings, revealing that a significantly higher proportion of children responded positively to the

multinutrient treatment compared to the placebo, as assessed by blinded clinician ratings using the CGI-I scale. This positive trend continued after transitioning from a blinded to an open-label RCT, where the placebo group also received multinutrient treatment (Leung et al., 2024). In this phase, the response rate for participants initially in the placebo group increased dramatically, rising by a factor of 2.78 (23% to 64%). The multinutrient group also experienced an increase in response rate from 53% to 66%. Additionally, while stimulant medication has been associated with height suppression (Faraone et al., 2008), multinutrient treatment has been shown to result in a greater height improvement (Johnstone et al., 2022; Leung et al., 2024). While evidence from the MADDY study highlights the potential benefits of multinutrient treatment alongside concerns about medication risks, these findings underscore the importance of accurate ADHD diagnosis and screening to ensure appropriate application and optimal outcomes in pediatric research.

1.4 ADHD Diagnosis and Screening in Pediatric Research

ADHD diagnosis requires signs of inattention and/or hyperactivity-impulsivity to be evident in multiple settings (such as at home, school, or work; with friends or relatives; in other activities), though children may present differently depending on the situation (American Psychiatric Association, 2022). The symptoms must not be explained by another mental, behavioral, or developmental disorder, and they should not be caused by a substance or medication (WHO, 2019). There is no single diagnostic test to determine whether a child has ADHD. Past studies have used a wide range of tools for ADHD diagnosis in children and youth including parental ratings, teacher ratings, self-reports, clinician interview, neuropsychological tests, biospecimen, electroencephalogram (EEG), and neuroimaging (Peterson et al., 2024a). Though various tools have been explored, a reliable diagnosis for ADHD depends on the clinical

assessment of an experienced clinician with supported by history of symptoms and input from multiple informants across different settings, including parents, teachers, and, in some cases, the child themselves (De Los Reyes & Kazdin, 2005; Eom & Kim, 2023). That said, the use of screening tools to gather data from multiple informants is a standard practice in pediatric studies, using validated tools (Martel et al., 2015; Mulraney et al., 2022; Staff et al., 2021). While collecting data from multiple informants offers a more comprehensive view of ADHD symptoms, it also introduces the challenge of inconsistent agreement across raters and measurement items. Together, these issues highlight the need to examine informant discrepancies and the factors contributing to such variation in ratings.

1.5 Informant Discrepancies and Multi-Informant Perspectives

According to the American Psychiatric Association (2013), ADHD symptom diagnosis should be made by collecting information from multiple informants. As such, ADHD screening commonly uses symptom rating scales completed by parents, teachers, and/or children. These screening measures are simple to use, can be efficiently administered to large populations, and provide a more comprehensive understanding of the child's condition by incorporating diverse perspectives from multiple informants based on their unique experiences with the child. However, there is evidence of varying degrees of agreement between different informants in ADHD populations (Antrop et al., 2002; Kemper et al., 2003; Mitsis et al., 2000; Murray et al., 2018; Sollie et al., 2013; Wolraich et al., 2004).

These differences may occur because raters see the child’s behavior in different settings and have their own viewpoints based on their relationship with the child. For example, parents see their child daily in various situations, while teachers mostly observe behavior in the classroom. Conversely, children may not fully understand their symptoms, which can affect how accurately

they report their own behavior (Kerr & Schneider, 2008). Moreover, the knowledge and expectation of the child normal development may influence the informant rating (Petric and Szamosközi, 2018). Discrepancies in rating can also be influenced by child characteristics (age, gender, race/ethnicity, social desirability, problem type), parent characteristics (anxiety, depression, stress, socio economic status), and family characteristics (family status, sibling birth order, number of siblings in the family, parental acceptance, level of parent-child interaction) (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005). In addition, rating discrepancies may also occur based on which parent provided the information and whether reports from both parents were combined (Duhig et al, 2000; Langberg et al., 2010; Sollie et al., 2013). For example, in a study of inter-parental agreement, fathers reported fewer problems than mothers and teachers on inattention in children. In analyses of sensitivity (the likelihood that a parent report confirms a positive teacher report) and specificity (the likelihood that a parent report confirms a negative teacher report), mother ratings showed higher sensitivity, whereas father ratings demonstrated greater specificity (Sollie et al., 2013). This suggests that mothers' ratings may be more sensitive for suspected ADHD in the child, while fathers' ratings better predict an ADHD diagnosis (Sollie et al., 2013). Other factors may be responsible for inconsistent agreement among raters, including difference in methods across studies (i.e., sample size and age range), informant perspective, child behavior across environment, and the different measures or scales used. Additionally, differences in data analysis strategies may also play a role in the mixed results.

1.6 Parent-Teacher Agreement in ADHD Symptoms and Related Disorders

In child psychopathology, parent-teacher ratings agreement estimate is generally low. For example, a meta-analysis of 119 studies reported an average parent-teacher correlation of 0.27 on

broadband measures of behavioral and emotional problems among children aged 1 to 19 years. The authors further noted that informants who interact with the child or adolescent in similar contexts (e.g., mothers and fathers) tend to show higher agreement in their reports than those who interact with the child or adolescent in different contexts (e.g., parents and teachers) (Achenbach et al., 1987). Specifically for children ADHD, previous studies have shown mixed results regarding parent-teacher agreement in identifying inattention and hyperactivity/impulsivity symptoms, with low to moderate effect sizes. While some research indicates greater agreement for inattention symptoms (Wolraich et al., 2004; Murray et al., 2018), other studies suggest higher agreement for hyperactivity/impulsivity symptoms (Mitsis et al., 2000; Narad et al., 2015; Sayal & Goodman, 2009; Sollie et al., 2013). In contrast, some research has identified no significant agreement between parent and teacher ratings of inattention and hyperactivity (Antrop et al., 2002). Discrepancies in the magnitude of ratings across informants have yielded inconclusive findings. Neither parents nor teachers consistently report higher or lower symptom severity than the other. While some studies indicate that teachers tend to rate symptoms as more severe than parents (Mitsis et al., 2000; Wolraich et al., 2004), others have found the opposite pattern in similarly aged samples, with teachers reporting lower symptom severity than parents (Antrop et al., 2002; Sollie et al., 2013). In addition, the level of informant rating agreement varies by samples characteristics such as age, sex, clinical vs non-clinical, overcontrolled vs under-controlled problems etc. (Achenbach et al., 1987). Moreover, parent-teacher ratings agreement is generally higher in clinically referred children compared to community samples (Achenbach et al., 1987; Sollie et al., 2013). Prior studies have also examined differences in agreement across behavioral domains, demonstrating that informant agreement is generally higher for externalizing behaviors (e.g., anger, aggression, hyperactivity)

than for internalizing behaviors (e.g., anxiety, depression) (Achenbach et al., 1987; Duhig et al., 2000; De Los Reyes & Kazdin, 2005). For example, Antrop et al. (2002) examined clinically diagnosed children with ADHD aged 6 to 12 years and found no significant parent-teacher agreement for inattention and hyperactivity/impulsivity symptoms; however, moderate to strong agreement was observed for symptoms related to ODD and CD. Similarly, a meta-analysis conducted on both clinical and non-clinical populations concluded that the overall effect size for parent-teacher agreement on ODD was moderate and significant (Petric & Szamosközi, 2018). Beyond core ADHD symptoms, researchers have also examined parent-teacher agreement in domains of emotional dysregulation, social and academic functioning, including peer interactions, compliance with instructions, classroom disruptive behaviors, assignment completion, and organizational skills. These studies consistently report low to moderate levels of agreement between informants (Mayes et al., 2016; McConaughy et al., 2011; Mitchison et al., 2022; Sollie et al., 2013; Tamm et al., 2021; Wolraich et al., 2004).

1.7 Parent-Child Agreement in Internalizing and Externalizing Symptoms

Children with ADHD are at greater risk of developing internalizing disorders like anxiety and depression with prevalence ranging from 12% to 50% for depression and from 15% to 35% for anxiety (Gnanavel et al., 2019). A meta-analysis conducted by Gershon & Gershon (2002) indicated that girls diagnosed with ADHD exhibit higher level of internalizing symptoms and are more likely to struggle with comorbid conditions like depression and anxiety. Children with ADHD and an internalizing disorder tend to experience more negative outcome than those with either disorder alone (Bryant et al., 2023). Despite the significance of these co-occurring disorders, research examining parent-child agreement on internalizing symptoms among children with ADHD showed mixed patterns. While parents and children both provide unique and

valuable information to the assessment of internalizing symptoms, agreement is generally low, particularly for less observable internalizing symptoms due to its subtle nature (Achenbach et al., 1987; Duhig et al, 2000; De Los Reyes & Kazdin, 2005). For example, Kemper et al. (2003) reported low mother-child rating concordance of internalizing symptoms in children aged 6-12 undergoing ADHD evaluations, although agreement improved when ratings focused on more observable behaviors such as sleep difficulties and fatigue rather than ideational symptoms like worry or difficulty making decisions. Parent-child agreement for internalizing symptoms tends to be higher for clinical samples than community samples but the direction of this discrepancy, whether child rated higher than parents, is inconsistent (Kemper et al., 2003). For example, children with clinical ADHD tend to self-report fewer depressive symptoms than their parents but the opposite pattern of discrepancy observed in general population (Frasher et al., 2018). Parent-child differences in perceived emotional and behavioral issues have also been observed in the assessment of anger and aggression in children with ADHD. For example, mothers reported increased aggression in ADHD children, whereas children self-reported more anger compared to non-ADHD children (Kitchens et al., 1999). Children with ADHD often experience peer problems, including rejection and victimization, and tend to have lower friendship quality and sociometric status than their peers without ADHD. In addition, these difficulties in peer relationships can contribute to negative consequences such as emotional distress, academic challenges, and long-term behavioral issues (Binti et al., 2022). Parent-child rating discrepancies have also been observed in the assessment of children's social functioning and peer relationships indicating low concordance (Limber et al., 2011; Varni et al., 2015). Sleep problems are common among children with ADHD. A systematic review reported an increased prevalence of sleep disturbances based on parental reports in both medicated and unmedicated children with ADHD

(Cohen-Zion & Ancoli-Israel, 2004) as well as greater ADHD symptom severity has been reported as a risk factor for persistent sleep issues (Lycett et al., 2014). Previous research has also shown that agreement between parent and child ratings of sleep disturbances in children with ADHD tends to range from low to moderate (Owens et al., 2000a; Owens et al., 2009).

1.8 Difference in Informant Agreement by Child Gender or Sex

It is well documented that ADHD and related emotional and behavioral symptom presentations differ by child gender, leading to disparities in referral patterns, diagnostic rates, and medication use (Ayano et al., 2023; Gershon & Gershon, 2002; Quinn & Madhoo, 2014; Rucklidge 2010). Empirical research has examined the association between child gender and informant agreement, with inconclusive findings (Achenbach et al., 1987). Studies suggest that, overall, rating discrepancies between informants may not be related to a child's gender; however, meaningful associations may appear within specific populations, such as clinical, community, or school-based groups (De Los Reyes & Kazdin, 2005). That said, in a meta-analytic review of behavioral and emotional problems among children aged 1 to 19 years, Achenbach et al., (1987) found no significant differences in informant rating correlations by child gender. Similarly, a meta-analysis by Duhig et al. (2000) examining inter-parental agreement reported no significant gender-related differences in mother-father rating correlations and discrepancies in emotional and behavioral problems in children and adolescents. However, comparatively little attention has been given to gender or sex differences in informant rating agreement specifically for the ADHD population.

1.9 Study Rationale

Most previous research on parent-teacher and parent-child ratings agreement have focused on individual symptom domains, such as ADHD core symptoms (Kennerley et al., 2018; Mitsis et al., 2000; Murray et al., 2007; Murray et al., 2018; Narad et al., 2015; Sayal & Goodman, 2009; Wolraich et al., 2004), ADHD and externalizing/internalizing problems (Antrop et al., 2002; Sollie et al., 2013), externalizing and internalizing problems (Frasher et al., 2018; Kemper et al., 2003; Kitchens et al., 1999; Petric & Szamoskozi, 2018; Varni et al., 2015), social or academic impairment (McConaughy et al., 2011; Tamm et al., 2021), or sleep disturbances (Owens et al., 2000a; Ownes et al. 2000b; Owens et al., 2009), without examining these co-occurring symptoms collectively. Furthermore, relatively little attention has been given to the differences in agreement by child sex. To the best of my knowledge, this would be the first study to investigate the parent-other adult and parent-child ratings agreement across a broad spectrum of ADHD and co-occurring emotional and behavior symptoms within a multisite randomized controlled trial (RCT) sample, while also assessing whether this agreement is moderated by child sex.

The rationale for examining these ratings agreement lies in the importance of understanding how different raters perceive and report ADHD and related symptoms. It will also underscore how differently a child may behave depending on the structure of the settings. ADHD is often accompanied by other emotional and behavioral disorders (Dickstein, 2016; Gnanavel et al., 2019; Lenzi et al., 2018), making it important to assess symptoms ratings agreement beyond ADHD core domains. Accurate assessment of these symptoms is crucial for diagnosis, treatment planning, and intervention strategies. As parent, teacher and child ratings are key sources of information, discrepancies in symptom ratings can pose significant challenges in clinical and research settings.

Another critical aspect of this investigation is the stratification of analyses by child sex. Evidence suggests that boys and girls may experience ADHD differently, with variations in symptom presentation, severity, and co-occurring conditions (Ayano et al., 2023; Gershon &Gershon, 2002; Gilbert et al., 2025; Quinn & Madhoo, 2014; Rucklidge, 2010), leading to differences in how their symptoms are rated across informants. Discrepancies in symptom ratings may contribute to underdiagnosis or misdiagnosis in certain groups. Stratified analysis will provide a deeper understanding of these sex-based differences, improving sex specific screening and assessment strategies in research settings.

It is hypothesized that informant agreement will be low to moderate for core ADHD symptoms as well as associated emotional and behavioral disorders. Agreement is expected to be higher for externalizing symptoms in boys and for internalizing symptoms in girls. Additionally, it is hypothesized that parents will report greater symptom severity across all domains compared to other informants.

1.10 Research Objective

The objective of this study is to assess the level of agreement of parent-other adult and parent-child report ratings on ADHD core symptoms and cooccurring emotional and behavior symptoms among children aged 6 - 12 years with ADHD and irritability using data from the MADDY study, and to investigate whether the ratings agreement have sex specific differences.

1.11 Research Questions

This study will examine the following research questions:

1. What is the level of agreement in parent vs other adult reported ratings on inattention, hyperactivity, social and academic impairment, ODD, DMDD, and anger issues of children with ADHD and irritability using CASI-5?
2. What is the level of agreement in parent vs child reported ratings on anger, anxiety, depression, peer relationships, and sleep issues using PROMIS?
3. Is the rating agreement moderated by child sex on both CASI-5 and PROMIS domains?

CHAPTER 2: METHODS

2.1 Study Design

This study is a secondary data analysis of two psychometric measures used in the MADDY study. The MADDY study was an eight-week randomized, multi-site, fully blinded, placebo-controlled trial that compared the efficacy and safety of a multinutrient with placebo (Johnstone et al., 2019). Randomized controlled trials (RCTs) are considered the gold standard in research for experimental study design or clinical trials. It minimizes bias and is a robust method to explore cause-and-effect relationships between an intervention and its outcome unlike other designs (Hariton & Locascio, 2018). The MADDY study received approval from the Institutional Review Boards at Oregon Health & Science University (OHSU) and The Ohio State University (OSU), as well as from the Conjoint Health Research Ethics Board at the University of Calgary, the US Food and Drug Administration, and Health Canada. It was also prospectively registered with the National Clinical Trials Registry.

2.2 Setting

The MADDY study collected data at three sites: Oregon Health & Science University (OHSU), Portland, Oregon; Ohio State University (OSU), Columbus, Ohio; and University of Lethbridge (UofL), Lethbridge, Alberta, Canada. A total of 135 participants were recruited through children's hospital, referrals from pediatricians and mental health providers, social media websites (e.g. Facebook), and school districts (Peachjar.com). Recruitment strategies varied by site.

2.3 Participants

2.3.1 Eligibility Criteria

In the MADDY study, the following eligibility criteria were used to recruit participants.

1) Be between 6 and 12 years old.
2) Be willing and able to swallow 9 to 12 capsules per day and attend all study appointments.
3) Meet DSM-5 criteria for ADHD, as determined by parent reports on the CASI-5 questionnaire, with symptoms present in multiple settings and impairing social or academic functioning.
4) Exhibit at least one symptom of irritability or anger based on the Oppositional Defiant Disorder (ODD) or Disruptive Mood Dysregulation Disorder (DMDD) subscale questions.
5) Be medication-free or washed out for at least two weeks before the in-person baseline assessment.
6) Be willing to provide blood samples at two time points: baseline and week 8, with an optional blood draw at week 16.

Concomitant treatments, including nutritional supplements, were allowed if they do not contain ingredients already included in the treatment product. Additional therapies, such as occupational therapy or counseling, were permitted. Information on concomitant treatments were collected and documented at each visit. Participants who report difficulty swallowing capsules first completed a pill-swallowing program developed by Kaplan et al., 2010 before the initial

study visit. Children who remain unable to swallow the capsules after completing this training were excluded from the study.

The MADDY study exclusion criteria, based on parent report, were as followed:

1) Presence of a neurological disorder affecting the brain or central functions (e.g., history or suspected intellectual disability, autism spectrum disorder, epilepsy, narcolepsy) or any major psychiatric condition requiring hospitalization (e.g., significant mood disorder, active suicidal ideation, or psychosis).
2) Presence of any serious medical condition, such as inflammatory bowel disease, a history of cancer, kidney or liver disease, hyperthyroidism, or diabetes.
3) Known allergy to any component in the intervention.
4) Known abnormalities in mineral metabolism (e.g., Wilson's disease, hemochromatosis).
5) Current use of any medication primarily affecting the central nervous system, including stimulants, or use of such medication within two weeks before enrollment.
6) Severe anxiety that prevents separation from a parent for study questionnaire completion.
7) Any disability that would interfere with the participant's ability to verbally answer questions.
8) Inability to communicate in English.
9) Pregnancy or being sexually active at baseline (applicable to girls who have started menstruating).

2.3.2 Source and Selection Method

In the MADDY study, each potential participant first completed a phone screening, where the study's purpose, procedures, risks, and benefits were explained. This screening also assessed

willingness to participate and determine eligibility. If a participant appears eligible based on the parent's responses, the parent received a link via email to complete the Child and Adolescent Symptom Inventory-5 (CASI-5) (Gadow & Sprafkin, 2013) eligibility questions through the Research Electronic Data Capture (REDCap) system (Harris et al., 2009). The eligibility questions included the ADHD, oppositional defiant disorder (ODD), and disruptive mood dysregulation disorder (DMDD) (American Psychiatric Association, 2013) symptom and impairment subscales. To qualify, children were assessed to meet a clinical threshold of at least six inattentive or hyperactive/impulsive symptoms from the ADHD subscale and exhibit at least one symptom of irritability or anger from the ODD or DMDD subscales. Severity was scored with 1 point for "often" or "very often" responses and 0.5 points for "sometimes" responses.

Phone screen data, including inclusion/exclusion criteria and CASI-5 responses, were reviewed weekly by the study Principal Investigators and research assistants at each site. They consulted a physician if needed on physical health-related eligibility questions. Challenging eligibility cases were discussed in weekly cross-site teleconferences to reach a consensus. Eligible participants were then scheduled for a baseline visit, where parental consent and child assent were obtained. The child received visual and verbal information about the study, using language suitable for a child as young as six years old.

2.4 Data Collection

In the MADDY study, recruitment and data collection occurred from April 2018 to June 2020. Following the signing of consent and assent forms, the participant underwent a fasting blood draw. Both the parent and child then completed baseline questionnaires using the web-based data management platform REDCap. Some questionnaires are self-reported, with

participants entering their responses directly into REDCap, while others are administered and recorded by study staff.

The parent provided the email address of a coach, teacher, or other adult familiar with the child to serve as a secondary reporter. This individual was asked to report on the child's ADHD, ODD, DMDD, and peer conflict symptoms from the CASI-5 questionnaire at baseline, week 8, and week 16. To ensure sufficient time for completing repeated assessments, teachers served as informants only if the child was enrolled during the first half of the school year. At the two U.S. sites, child and parent participants received compensation for attending each visit, with additional compensation provided to children for each biological sample collected. In Alberta, Canada, where offering incentives to study participants is not customary, parents were instead be reimbursed for travel expenses, including parking and mileage, for each onsite visit.

2.5 Sample Size

In the MADDY study, the intention-to-treat (ITT) sample consists of 126 participants. As the present study is a secondary analysis, we used the same sample to examine the above-mentioned research questions.

2.6 Measures

2.6.1 Child & Adolescent Symptom Inventory–5 (CASI-5)

The Child & Adolescent Symptom Inventory–5 (CASI-5) is widely used in research to assess DSM-5 emotional and behavioral disorder symptoms in children and adolescents aged 5 to 18 years. It is derived from the CASI-4R (Gadow & Sprafkin, 2005), which combined items from the Child Symptom Inventory (CSI) for 5–12-year-olds in elementary school and the Adolescent Symptom Inventory (ASI) for 12–18-year-olds in middle or secondary schools.

Empirical evidence indicates that the CASI-4R demonstrates strong psychometric properties, including high internal consistency, test–retest reliability, and convergent and discriminant validity across both community and clinical samples (Gadow & Sprafkin, 2015). In the teacher version of the CSI-4, Cronbach's alpha coefficients were high for ADHD Inattentive type (ADHD-I; α = .93), ADHD Hyperactive-Impulsive type (ADHD-HI; α = .93), ADHD Combined type (ADHD-C; α = .93), and Oppositional Defiant Disorder (ODD; α = .92). Reliability was moderate to strong for Conduct Disorder (CD; α = .78), Generalized Anxiety Disorder (GAD; α = .76), and Social Phobia (α = .72) (Gadow et al., 2004). Similarly, the parent version of the CSI-4 showed strong reliability for both Symptom Severity and Symptom Count scores. The respective alpha values were ADHD-I (.92, .88), ADHD-HI (.91, .86), ADHD-C (.91, .87), ODD (.91, .86), CD (.79, .73), GAD (.75, .69), and Social Phobia (.77, .37) (Sprafkin et al., 2002).

In the present study, a 43-item rating scale from CASI-5 was used to evaluate ADHD symptoms (inattention 9 items; hyperactivity 10 items), ODD (8 items) and DMDD (2 items) symptoms, anger and aggression towards peers (10 items) and impairment (4 items) for each syndrome. Items were rated from 0 (never) to 3 (very often). The final score of each symptom was calculated by finding the mean of corresponding item ratings. Both parents and other adults such as teacher, grandparent, aunt/uncle, or neighbor who knew the child's behavior well completed the CASI-5 at baseline and week 8.

2.6.2 Patient-Reported Outcomes Measurement Information System (PROMIS)

The Patient-Reported Outcomes Measurement Information System (PROMIS) is a comprehensive set of self-report measures designed to assess physical, mental, and social health across diverse populations, including children and adolescents. It was developed to improve the assessment of patient-reported outcomes (PROs) and using item response theory (IRT), items are

calibrated to produce scales that aim to maximize reliability and validity across the full range of the underlying latent constructs (Varni et al, 2014; Varni et al., 2015). PROMIS provides reliable and precise measurement of health domains such as anxiety, depression, anger, fatigue, pain, sleep disturbance, physical functioning, and peer relationships. It has been widely used in both clinical and research settings, with strong evidence supporting its precision, reliability and validity across various health domains (Cella et al., 2010; Varni et al., 2014).

Cella et al. (2010) reported that, the PROMIS adult self-reported Sleep Disturbance, Anger, Anxiety, and Depression banks demonstrate excellent psychometric properties. Across domains, short forms show very high correlations with their respective full item banks ($r = 0.96$), indicating that the shorter versions accurately represent the full measures. The Sleep Disturbance full bank demonstrates strong reliability (> 0.88) and convergent validity with the Pittsburgh Sleep Quality Index ($r = 0.85$). The Anger bank similarly demonstrated excellent reliability, exceeding 0.93 with moderately correlated with the Aggression Questionnaire ($r=0.51$). The Anxiety bank showed reliability greater than 0.89. It demonstrates strong convergent validity with established measures, including the Mood and Anxiety Symptom Questionnaire ($r = 0.80$) and the Center for Epidemiological Studies-Depression Scale ($r = 0.75$). Similarly, the Depression bank showed reliability greater than 0.92. The depression bank also demonstrates strong correlations with legacy measures (Center for Epidemiological Studies-Depression Scale $r = 0.83$; Mood and Anxiety Symptom Questionnaire $r = 0.72$), further supporting construct validity.

Varni et al. (2014) examined the psychometric properties of the PROMIS pediatric scales and found strong evidence of reliability and score comparability. Scores were highly consistent across different scoring and administration methods, including computer adaptive testing and

static short forms, as well as response pattern scoring versus summed score conversion. The study reported high internal consistency across domains, with coefficients exceeding 0.85 for Depressive Symptoms, 0.83 for Anxiety, 0.79 for Anger, and 0.83 for Peer Relationships. Test–retest reliability was also generally strong, with correlations ranging from 0.70 to 0.80 for Depressive Symptoms, Anxiety, and Peer Relationships. The Anger scale demonstrated slightly lower but still acceptable stability over time, with a test–retest correlation of approximately 0.60. Past research also demonstrated strong reliability (>0.90) of the PROMIS Pediatric Sleep Disturbance and Sleep-Related Impairment. Evidence for validity was supported by correlations with existing measures of pediatric sleep health for children with sleep problems and those with chronic and neurodevelopmental disorders (Forrest et al., 2018).

In this study, the PROMIS was used to assess anxiety, depression, peer relationships, anger, and sleep disturbance. Each domain includes between 4 and 13 items, rated on a 5-point Likert scale ranging from 1(Never) to 5(Always). Scores for each symptom were derived by averaging the responses across the relevant items. Both parents and children independently completed these scales at baseline and week 8.

2.6.3 Demographic Characteristics

Parents provided demographic information about their children including age (in years), biological sex (male; female), race (American Indian or Alaska native, Asian, African American/Black, Native Hawaiian or Pacific Islander, White, Other), ethnicity (Hispanic or Latino, Not Hispanic or Latino, Other). They also reported their marital status (Single, Married or Common-law, Divorced), education level (No education; Elementary; High School, Technical College, University or Higher), household income (USD ≤30k, 30k–60k, 60k–80k, 80k+), and the number of people in their household.

2.6.4 Clinical Characteristics

Additionally, parents reported the child's health-related behaviors, such as past ADHD medication history (Yes/No), physical and emotional health (Excellent; Very good; Good; Fair, Poor), antibiotic use (Yes/No) and their frequency (times per year; times since birth), asthma (Yes/No), learning problems (Yes/No), bedwetting (Yes/No), and their status (Past; Present; Both). They also provided information on their own mental health behaviors, including anxiety (Yes/No), depression (Yes/No), mania (Yes/No), psychotic disorder (Yes/No), learning problems (Yes/No), ADHD symptoms (Yes/No), behavioral issues (Yes/No), alcohol use (Yes/No), and legal or illegal drug use (Yes/No).

CHAPTER 3: ANALYSIS STRATEGY

3.1 Sample Description

Descriptive analyses were conducted for the demographic and clinical characteristics of the participating children. Continuous variable (e.g., child age) was summarized using mean and standard deviation, whereas categorical variables (e.g., race, ethnicity, medication history, comorbidities etc.) were summarized using frequencies and percentages. To account for potential sex-related differences, descriptive statistics were stratified by child sex and presented in tabular form.

Prior to inferential analyses, the distributions of continuous variables were examined. Normality of the data was evaluated using histograms, Q-Q plots, and boxplots, to assess whether the assumptions underlying parametric tests were met. Outliers were assessed through boxplots and scatterplots.

3.2 Agreement Assessment

To evaluate the degree of agreement between different raters (i.e., parent, other adult, and child reports), multiple statistical approaches were applied. First, the bivariate association between rater scores was examined using Pearson product–moment correlation coefficients. To test for differences between paired rater scores, dependent samples t-tests were conducted. While correlation coefficients provide insight into the linear association between scores, they do not quantify the absolute agreement; therefore, additional agreement indices were computed.

Initially, intraclass correlation coefficients (ICC) were calculated based on a linear mixed-effects model framework, with child included as a random factor and rater as a fixed factor. This approach allowed estimation of variance attributable to between-rater and within-rater sources.

Assumptions of the linear mixed-effects model were examined, with particular attention to the distribution of residuals. Inspection of residual plots and normality tests (e.g., Shapiro-Wilk) indicated that the assumption of normality was not satisfied for several symptom domains.

Given these violations, the analytic strategy was revised. Agreement was instead quantified using the Concordance Correlation Coefficient (CCC), which evaluates both precision and accuracy of agreement (Lin et al., 2002). The CCC integrates measures of correlation and systematic bias that refers to consistent differences in the level of scores between raters. Because the CCC is a sample estimate, a confidence interval is required to quantify the precision of the agreement estimate and to determine how much uncertainty surrounds the observed CCC value. The distribution of rating differences was non-normal, and the CCC does not have a closed-form standard error under such conditions. For this reason, confidence intervals were obtained using a nonparametric bootstrap procedure, which provides valid inference without relying on normality assumptions. In the bootstrap framework, 2000 bootstrap samples were generated by repeatedly drawing observations with replacement from the original dataset, each sample being the same size as the original. For each bootstrap sample, the CCC was computed, producing an empirical sampling distribution. 95% confidence intervals were then derived from this bootstrap distribution by ordering the resampled CCC estimates and identifying the 2.5th and 97.5th percentiles of the distribution. A CCC value ranges from -1 to +1, where +1 indicates a perfect agreement, -1 indicates a perfect disagreement, and 0 indicates no agreement (Lin et al., 2002). According to McBride's (2005) guideline, the strength of the agreement was categorized as: almost perfect (> 0.99), substantial (0.95-0.99), moderate (0.90-0.95), and poor (< 0.90).

Since raters provided symptom ratings at two time points (baseline and week 8), CCC values were computed separately for each time point. To further explore potential sex differences,

analyses were stratified by child sex, and CCC values were estimated for boys and girls independently. Comparisons of agreement across subgroups (i.e., baseline vs. week 8 and boys vs. girls) were tested using the bootstrap framework. Because the sampling distribution of the CCC is not analytically tractable, particularly under non-normal data and unequal subgroup sample sizes.

CCC provides a single overall index of agreement, however, it does not reveal how agreement varies across the range of symptom severity. Consequently, the Bland-Altman plot was used to strengthen the analysis. This method assesses individual-level agreement across the measurement range and identify both systematic bias and variability in rating differences (Bland and Altman, 1999). For each rater comparison, Bland-Altman plots were generated by plotting the difference between paired rater scores against their mean. The average scores of the two raters were plotted on the x-axis, while the difference between these scores were plotted on the y-axis. A horizontal line represents the mean of rating differences, and two additional lines mark the limits of agreement (mean of rating differences ±1.96 * standard deviations of the differences), which allows to determine whether most observations fell within the limits, as well as to evaluate potential systematic bias and heteroscedasticity across the measurement range. Agreement between raters was evaluated using Bland-Altman plots by examining the mean difference to assess systematic bias, the width of the 95% limits of agreement (LOA) to determine whether raters could be considered interchangeable, and proportional bias, assessed by regressing the rating differences on their mean. Raters were considered non-interchangeable when LOA were wide relative to the scale or when systematic or proportional bias was present. The percentage of observations falling outside the LOA was also reported to identify occasional large disagreements. In addition, to examine whether rating differences exhibited an upward or

downward trend with increasing mean scores, we conducted a linear regression with robust standard errors and tested whether the slope of the fitted line significantly differed from zero. The shaded region around the fitted line represents the uncertainty associated with the estimated regression slope (used to assess proportional bias), providing a visual indication of variability in the trend.

Statistical analyses were performed using Stata version 18 (StataCorp LLC, College Station, TX) and the R statistical environment. A two-tailed p-value < 0.05 was considered statistically significant.

3.3 Missing Data

Given the very low proportion of missing responses ($\leq 3\%$) across both time points and measures, the listwise deletion method was employed. At baseline, the CASI-5 questionnaire was completed by parent-other adult dyad for 108 children, with only one missing response each for DMDD, ODD, and anger. At week 8, 95 parent-other adult dyads completed the questionnaire, with a single missing response for impairment. For the PROMIS questionnaire, 126 child-parent dyads completed it separately at baseline, with one missing response for anger, four for peer relations, and three for sleep. At week 8, 120 child-parent dyads completed the questionnaire, with one missing response each for anger and depression, and two missing responses each for peer relations and sleep.

CHAPTER 4: RESULT

4.1 Participant Characteristics

Table 1 presents the sociodemographic and clinical characteristics of the study participants. The sample included 126 children, with a mean age of 9.8 years. Of those, 92 were male (73%) and 34 were female (27%). Most children were White (90%), and 72.5% were non-Hispanic or Latino. More than half of the participants (57.6%) were from households with annual incomes above $80,000, and 61.1% had parents with a university degree or higher. Nearly half (46.6%) had a history of medication use for ADHD or other psychological conditions, and 50% had documented learning difficulties. Among parents, 23.4% reported a history of learning problems, 35.5% reported behavioral problems, and 52.4% had ADHD.

4.2 Concordance Correlation Coefficients (CCCs) for Parent–Other Adult Reports on CASI-5 Domains

Baseline CCCs for all CASI-5 symptom domains indicated poor parent-other adult agreement ranging from 0.12 (95% CI: 0.00-0.24) for inattention to 0.43 (95% CI: 0.27-0.57) for disruptive mood dysregulation disorder (DMDD) (see Table 2). Among these, highest agreement was observed on impairment (0.43 (95% CI: 0.28-0.56)) and DMDD (0.43 (95% CI: 0.27-0.57)). At week 8, CCCs were generally higher, ranging from 0.34 (95% CI: 0.15-0.53) for DMDD to 0.53 (95% CI: 0.40-0.65) for hyperactivity, with CCC for inattention domain significantly higher at week 8 than baseline (difference in CCC=-0.27 (95% CI: -0.44 to -0.09)).

4.3 CCCs for Parent–Child Reports on PROMIS Domains

Table 3 summarizes the CCCs for parent-child agreement across the five PROMIS domains. At baseline, agreements were poor across all five domains with CCCs ranging from 0.07 (95%

CI: -0.13-0.27) for anxiety to 0.39 (95% CI: 0.24-0.54) for sleep. Except for anxiety, agreement at baseline was slightly higher than at week 8; however, the difference in CCCs between baseline and week 8 were not statistically significant.

4.4 Sex-Stratified Analyses

Table 4 presents CASI-5 sex-stratified results. At baseline, among boys, parent-other adult agreement was poor for all symptom domains, with the strongest agreement observed on impairment (0.46 (95% CI: 0.32-0.60)) and DMDD (0.46 (95% CI: 0.30-0.61)). Among girls, overall parent-other adult agreement was slightly lower than that was observed in boys for symptom domains other than core ADHD symptoms (inattention and hyperactivity). Within the girl subgroup, hyperactivity showed the highest level of agreement (0.44, 95% CI: 0.13-0.69) among the CASI-5 six domains. The differences in CCCs between boys and girls at baseline were not statistically significant. At week 8, parent-other adult agreement among boys remained poor across all six CASI-5 symptom domains. However, the magnitude of agreement was higher than baseline for all domains except impairment and DMDD. In contrast, week 8 agreement among girls improved across all six domains relative to baseline. Moreover, except for anger, agreement at week 8 was higher among girls than boys, with significantly greater agreement observed for inattention in girls compared with boys.

Table 5 presents PROMIS sex-stratified results. At baseline, parent-child agreement among boys was poor across domains, with relatively higher agreement in more observable domains (i.e., anger, peer relations, and sleep) than less observable domains (i.e., anxiety and depression). Among girls, the strongest agreement was observed for sleep (CCC = 0.53, 95% CI: 0.28 - 0.72), exceeding that of boys. A significant sex difference was observed only for peer relations (difference in CCC=0.46, (95% CI: 0.08 to 0.81)). However, the associated wide confidence

interval suggests substantial uncertainty in the magnitude of this difference, likely reflecting the smaller sample size among girls. Consistent with baseline findings, parent-child agreement among boys at week 8 remained poor, with relatively higher agreement for anger, peer relations, and sleep. Among girls, the highest level of agreement at week 8 was observed in peer relationship (CCC = 0.32, 95% CI: -0.07 - 0.67). No statistically significant sex differences were observed at week 8.

4.5 Summary of Informant Ratings

The mean and standard deviation were calculated for parent, other adult, and child self-report ratings at both baseline and week 8. Differences between raters (parent vs. other adult; parent vs. child) were then examined using paired-samples t-tests.

Table 6 presents the summary of parent and other adult ratings across six CASI-5 domains. At both baseline and week 8, inattention was rated highest by both parents and other adults, while anger issues received the lowest ratings. Mean scores from both raters across all six domains were lower at week 8 compared to baseline, indicating symptom improvement over time. Except for anger at week 8, across both time points, parents tended to provide higher ratings than other adults, with the average rating difference at baseline ranging from 0.03 for anger to 0.45 for inattention and at week 8 ranging from -0.08 for anger to 0.19 for ODD. At baseline, significant parent-other adult rating difference was observed in all domains except anger, while at week 8, significant rating difference persisted only for hyperactivity and ODD.

Table 7 summarizes parent and child ratings across five PROMIS domains. At both baseline and week 8, peer relations were rated highest by both parents and children, whereas depression received the lowest ratings. Mean score differences across domains between baseline and week 8

indicates an overall improvement in symptoms over time. Across both time points, parents rated anger symptoms significantly higher than children did. In contrast, children reported significantly greater sleep problems than their parents at week 8.

4.6 Summary of Informant Ratings by Child Sex

Sex-stratified comparisons of mean differences between raters (parent vs. other adult; parent vs. child) were evaluated using paired samples t-tests conducted separately for boys and girls at both baseline and week 8.

Table 8 summarizes parent and other adult ratings across the six CASI-5 domains stratified by child sex. At both baseline and week 8, inattention received the highest ratings from both parents and other adults for boys and girls, whereas anger problems received the lowest ratings. Mean scores across all domains decreased from baseline to week 8 for both raters and both sexes, reflecting overall symptom improvement over time.

At baseline, among boys, parents provided significantly higher ratings than other adults for all CASI-5 domains except anger. Among girls, parents rated only inattention significantly higher than other adults. By week 8, parent-other adult discrepancies shifted: for boys, parents rated ODD symptoms significantly higher than other adults, whereas other adults rated anger symptoms significantly higher than parents. Among girls, parents rated hyperactivity symptoms significantly higher than other adults.

Table 9 summarizes parent and child ratings across the five PROMIS domains stratified by child sex. At both baseline and week 8, peer relations received the highest ratings from both parents and children for boys and girls, while depression received the lowest ratings. The

difference in mean symptoms levels rated by parent and child across domains between baseline and week 8 indicates an overall improvement in symptoms over time.

Across both time points, parents rated their child's anger symptoms significantly higher than the child's self-ratings for both boys and girls. At baseline, among boys, parents also rated peer relationship significantly higher than the boys' own reports.

4.7 Bland-Altman Plots

Bland-Altman plots are widely used to assess agreement between two raters because they provide both a visual and quantitative way to examine bias and variability. By plotting the differences between ratings against their mean, the plot makes it possible to detect systematic bias, reflected as a consistent deviation of the average difference from zero. It also helps identify proportional bias, which occurs when the differences increase or decrease across the range of ratings. In addition, Bland-Altman analysis estimates the limits of agreement, which represent the interval within which most differences between the two ratings lie. These limits are crucial for determining whether the limits of agreement are narrow enough to indicate acceptable agreement or interchangeability between the two raters.

4.7.1 Parent-Other Adult Agreement (CASI-5)

Bland-Altman plot analyses were conducted to examine parent-other adult agreement across CASI-5 domains at baseline and week 8. At baseline, the mean rating differences ranged from 0.03 to 0.45 (see Table 6), indicating that parents tended to rate symptoms higher than other adults on average. At week 8, the mean rating differences across domains were smaller, ranging from -0.08 to 0.19 (see Table 6), suggesting improved parent-other adult agreement. These findings align with the CCC values presented in Table 2, which also showed higher concordance

at week 8 compared to baseline. At baseline, the width of 95% limits of agreement (LOA) ranged from 2.6-3.5 points on the 0-3 scale, suggesting the two informants are not interchangeable. Similarly, at week 8, the LOA widths narrowed slightly ranging from 2.0-3.3 points, suggesting limited interchangeability between informants despite modest improvement in agreement over time.

For inattention, at baseline, parents rated children on average 0.45 points higher than other adults (LOA: -0.95 to 1.85). A significant downward trend (p-value=0.008) was observed in difference of symptom ratings with its increasing mean, such that parent-other adult rating discrepancy were higher at low-to-moderate symptom severity, with gap narrowing at higher severity. No strong visual evidence of increasing variability across the range of means, although a small number (4.6%) of observations fell outside the limits of agreement reflected occasional large disagreements. On the other hand, at week 8, mean rating difference was smaller (Diff =0.12 points) than baseline. That is, parents rated, on average, 0.12 points higher than other adults, with narrowed LOA (LOA: -1.16 to 1.40). The findings also align with the CCC values which showed parent-other adult agreement improved significantly at week 8 than baseline (difference in CCC= - 0.27, CI: -0.44 to -0.09). At week 8, no systematic trend or strong visual evidence of increasing variability in the differences across the range of means was observed, with only 3.2% of observations falling outside the LOA.

For hyperactivity, at baseline, the mean rating difference was 0.25 points (LOA: -1.46 to 1.95), with a significant downward trend (p-value=0.001) indicating proportional bias, i.e., parent-other adult rating differences were higher at low symptom severity and lower at higher severity. No strong pattern of variability in differences was observed, with only 4.6% of observations outside the limits. At week 8, the mean rating difference was again smaller (Diff

=0.15 points) with narrowed LOA (LOA: -1.25 to 1.53) indicating higher agreement at week 8 than baseline aligning with CCC values. At week 8, no significant systematic trend was found and about 5.3% of observations were outside the LOA.

For impairment, at baseline, the mean rating difference was 0.28 points (LOA: -1.23 to 1.78), again showing, on average, parents rated symptoms higher than other adults. The significant downward trend (p-value=0.002) was evident but no substantial visual evidence of increasing variance in differences, and 6% of observations falling outside the limits. At week 8, agreement was stronger than baseline, aligning with CCC values, indicating parents rated only 0.06 points higher on average (LOA: -1.24 to 1.37). Notably, at week 8, a significant downward trend (p-value=0.003) was apparent indicating more positive differences (parents rated higher than other adults) at lower mean severity, and the differences decreases or become negative (other adults rated higher than parents) at higher mean severity, on average. About 5.3% of observations falling outside the LOA.

For ODD, at baseline, parents rated 0.39 points higher on average (LOA: -1.29 to 2.06) than other adults. The scatterplot showed a significant downward pattern (p-value=0.002) indicating parent-other adult rating differences get lower as severity increases. About 4.6% of observations outside the limits. At week 8, the mean rating difference was lower to 0.19 (LOA: -1.28 to 1.66), with a downward trend (p-value=0.003) indicating higher rating differences at low and high symptoms but lower at moderate symptoms. About 3.2% of observations outside the LOA.

For DMDD, at baseline, the mean rating difference was 0.29 points (LOA: -1.44 to 2.03), with no strong visual pattern of variance in differences. No significant trend in differences was observed. About 6.5% of observations fell outside the LOA. At week 8, parents rated on average

0.15 points higher than other adults (LOA: -1.51 to 1.80). Again, showing no significant trend, with 5.3% of observations outside the LOA.

Finally, for anger, at baseline, parent and other adult ratings were nearly identical (Diff = 0.03, LOA: -1.26 to 1.32). The scatterplot showed a funnel shape, suggesting proportional bias, i.e., lower discrepancies at lower levels of anger but higher discrepancies at higher severity. About 5.6% of observations were outside the limits indicating occasional large disagreements. At week 8, the mean rating difference was negative (-0.08), indicating other adults rated anger slightly higher than parents on average (LOA: -1.04 to 0.87). The scatterplot again showed a funnel pattern, suggesting lower discrepancies at lower levels with higher discrepancies at higher levels. About 7.4% of observations fell outside the LOA.

4.7.2 Parent-Child Agreement (PROMIS)

Bland-Altman plot analyses were also conducted to examine parent-child agreement across PROMIS domains at baseline and week 8. The analyses revealed mean parent-child rating differences at baseline ranging from -0.14 to 0.43 (see Table 7) while these were slightly larger at week 8 ranging from -0.17 to 0.50 (see Table 7). These findings are consistent with CCC values in Table 3, which showed higher concordance at baseline compared to week 8. At baseline, the width of limits of agreement ranging from 3.3 to 4.2, underscoring limited interchangeability between raters. Similarly, at week 8, LOA widths ranging from 3.5 to 4.3, again showing child and parent ratings are not interchangeable.

For anger, at baseline, a positive mean rating difference of 0.43 suggested parent rated symptoms higher than children on average, with limits of agreement ranging from -1.64 to 2.51. A significant upward trend (p-value=0.038) was evident, indicating smaller discrepancies at

lower levels and discrepancy increased as symptoms severity increased. About 4% of observations fell outside the LOA. At week 8, the mean rating difference remained positive (Diff =0.50; LOA: -1.34 to 2.35), again reflecting higher ratings by parents. A significant upward trend (p-value=0.004) and funnel-shaped pattern was evident suggesting smaller discrepancies at lower severity but higher discrepancies at higher severity. About 5.9% of observations fell outside the LOA.

For anxiety, at baseline, parents rated slightly lower than children on average (Diff = - 0.14, LOA: -2.01 to 1.73). The scatterplot showed a funnel shape, with smaller discrepancies at lower levels with higher discrepancies at higher severity. About 4.8% of observations lie outside of LOA. At week 8, parents continued to rate slightly lower (Diff = - 0.09, LOA: -1.83 to 1.65), with a similar funnel pattern reflecting greater discrepancies at higher symptom severity. About 5% of observations lie outside the LOA.

For depression at baseline, mean rating difference was minimal (Diff =0.04), with limits of agreement -1.83 to 1.90. A similar funnel pattern was again observed indicating greater discrepancies at higher severity levels, with 4.8% of observations outside the LOA. At week 8, mean rating difference remained minimal (Diff =0.04, LOA: -1.75 to 1.84). A significant upward trend (p-value=0.029) and funnel pattern was again observed, with 7.6% of observations outside the limits.

For peer relations at baseline, the mean rating difference was 0.07 (LOA: -1.75 to 1.88), indicating near equivalence between raters. A reverse funnel shape pattern was observed suggesting smaller discrepancies at higher levels of symptom severity but higher discrepancies at lower levels of symptoms. About 4.1% of observations lie outside the LOA. At week 8, mean rating difference was slightly higher to 0.16 (LOA: -1.98 to 2.30) indicating higher average

discrepancy at week 8 than baseline which also aligns with CCC values in Table 3. A significant downward trend (p-value=0.01) in difference in symptom ratings was evident. A reverse funnel trend suggested smaller discrepancies at moderate to severe level of symptoms. About 6.8% of observations fell outside the LOA.

For sleep disturbance at baseline, the negative mean rating difference was - 0.07 (LOA: -1.75 to 1.61) indicating, on average, parents rated 0.07 points lower than children. A significant upward trend (p-value=0.021) suggesting larger discrepancies at both low and high severity levels but smaller rating differences at moderate levels. About 4.9% of observations outside the LOA. At week 8, parents continued to rate sleep issues lower than their child (Diff = -0.17, LOA: -1.92 to 1.58). No significant directional trend was observed, though a funnel shape indicated lower discrepancies at lower severity levels. About 5.1% of observations fall outside the LOA reflecting occasional large disagreements.

Table 1 – Socio Demographics and Clinical Characteristics of Study Participants

Characteristics	Total (n=126)	Boy (n=92)	Girl (n=34)
Child age, mean (SD)	9.8 (1.7)	9.8 (1.7)	9.9 (1.7)
Race[a]			
Asian	3 (2.7)	2 (2.4)	1 (3.4)
African American/Black	4 (3.6)	2 (2.4)	2 (6.9)
White	100 (90.1)	76 (92.8)	24 (82.8)
Other[b]	4 (3.6)	2 (2.4)	2 (6.9)
Ethnicity[a]			
Not Hispanic or Latino	71 (72.5)	50 (70.4)	21 (77.8)
Hispanic or Latino	6 (6.1)	5 (7.0)	1 (3.7)
Other[b]	21 (21.4)	16 (22.6)	5 (18.5)
Parent marital status			
Married or common law	100 (79.4)	72 (78.3)	28 (82.4)
Divorced	15 (11.9)	12 (13.0)	3 (8.8)
Single	11 (8.7)	8 (8.7)	3 (8.8)
Parent education Level			
High school	20 (15.9)	11 (12.0)	9 (26.5)

Technical college/trade school	29 (23.0)	23 (25.0)	6 (17.6)
University or higher	77 (61.1)	58 (63.0)	19 (55.9)
Family household income[a]			
≤ $30,000	12 (9.6)	8 (8.8)	4 (11.8)
> $30,000 to ≤ $60,000	27 (21.6)	21 (23.1)	6 (17.7)
> $60,000 to ≤ $80,000	14 (11.2)	10 (11.0)	4 (11.8)
> $80,000	72 (57.6)	52 (57.1)	20 (58.8)
Child health issues			
Medication History for ADHD or other psychological issues[a]			
Yes	55 (46.6)	41 (48.2)	14 (42.4)
No	63 (53.4)	44 (51.8)	19 (57.6)
Asthma			
Yes	15 (11.9)	11 (12.0)	4 (11.8)
No	111 (88.1)	81 (88.0)	30 (88.2)
Learning issue			
Yes	63 (50.0)	43 (46.7)	20 (58.8)
No	63 (50.0)	49 (53.3)	14 (41.2)
Bed wetting			
Yes	20 (15.9)	15 (16.3)	5 (14.7)
No	106 (84.1)	77 (83.7)	29 (85.3)
Parents' health issues			
Anxiety status[a]			
Yes	45 (36.3)	36 (40.0)	9 (26.5)
No	79 (63.7)	54 (60.0)	25 (73.4)
Major depression[a]			
Yes	22 (17.7)	18 (20.0)	4 (11.8)
No	102 (82.3)	72 (80.0)	30 (88.2)
Manic depression[a]			
Yes	8 (6.4)	7 (7.8)	1 (2.9)
No	116 (93.6)	83 (92.2)	33 (97.1)
Psychotic disorder[a]			
Yes	1 (0.8)	0 (0.0)	1 (2.9)
No	123 (99.2)	90 (100.0)	33 (97.1)
Learning problems[a]			
Yes	29 (23.4)	22 (24.4)	7 (20.6)
No	95 (76.6)	68 (75.6))	27 (79.4)
ADHD[a]			
Yes	65 (52.4)	45 (50.0)	20 (58.8)
No	59 (47.6)	45 (50.0)	14 (41.2)
Behavior problem[a]			
Yes	44 (35.5)	33 (36.7)	11 (32.4)
No	80 (64.5)	57 (63.3)	23 (67.7)
Alcohol consumption[a]			

Yes	16 (12.8)	12 (13.2)	4 (11.8)
No	109 (87.2)	79 (86.8)	30 (88.2)
Legal drug use[a]			
Yes	6 (4.8)	3 (3.3)	3 (8.8)
No	118 (95.2)	87 (96.7)	31 (91.2)
Illegal drug use[a]			
Yes	10 (8.0)	5 (5.5)	5 (14.7)
No	115 (92.0)	86 (94.5)	29 (85.3)

Note: ADHD = attention-deficit/hyperactivity disorder.

[a]Fifteen missing race, 28 missing ethnicity, 1 missing income, 8 missing child medication history, 2 missing parent health condition, 1 missing alcohol and illegal drugs, and 2 missing legal drug use.

[b]Includes American Indian/Native American or Alaska Native, Metis, Native Hawaiian or Other Pacific Islander or other for race; Jewish, Japanese, or other for ethnicity.

Table 2: CASI-5 Parent-Other Adult Agreement

Characteristics	Baseline				Week 8				Diff in CCC (Baseline – Week 8)
	n	r	C.b	CCC (95% CI)	n	r	C.b	CCC (95% CI)	
Inattention	108	0.17	0.71	**0.12 (0.00, 0.24)**	95	0.40**	0.98	**0.39 (0.24, 0.52)**	**-0.27 (-0.44, -0.09)**
Hyperactivity	108	0.40**	0.90	**0.36 (0.20, 0.51)**	95	0.55**	0.96	**0.53 (0.40, 0.65)**	-0.17 (-0.33, 0.01)
Impairment	108	0.48**	0.90	**0.43 (0.28, 0.56)**	94	0.51**	0.92	**0.47 (0.29, 0.63)**	-0.04 (-0.22, 0.14)
ODD	107	0.44**	0.89	**0.39 (0.26, 0.52)**	95	0.52**	0.92	**0.48 (0.34, 0.61)**	-0.09 (-0.24, 0.06)
DMDD	107	0.46**	0.93	**0.43 (0.27, 0.57)**	95	0.35**	0.97	**0.34 (0.15, 0.53)**	0.11 (-0.09, 0.27)
Anger	107	0.34**	0.99	**0.34 (0.16, 0.52)**	95	0.48**	0.96	**0.46 (0.27, 0.62)**	-0.12 (-0.34, 0.10)

Note: r= Pearson correlation coefficient; C.b=Bias correction factor; CCC=Bootstrapped Concordance correlation coefficient; CI: Confidence interval; *p-value < 0.05 and **p-value < 0.01

Table 3: PROMIS Parent-Child Agreement

Characteristics	Baseline				Week 8				Diff in CCC (Baseline – Week 8)
	n	r	C.b	CCC (95% CI)	n	r	C.b	CCC (95% CI)	
Anger	125	0.28**	0.86	**0.24 (0.09, 0.37)**	119	0.29**	0.79	**0.23 (0.09, 0.39)**	0.01 (-0.18, 0.18)
Anxiety	126	0.08	0.88	0.07 (-0.13, 0.27)	120	0.19*	0.95	**0.18 (0.01, 0.35)**	-0.11 (-0.32, 0.10)
Depression	126	0.18*	0.94	0.17 (-0.03, 0.37)	119	0.09	0.99	0.09 (-0.12, 0.31)	0.08 (-0.09, 0.26)
Peer Relation	122	0.37**	0.99	**0.37 (0.25, 0.50)**	118	0.25**	0.99	**0.25 (0.08, 0.42)**	0.13 (-0.07, 0.34)
Sleep	123	0.39**	0.99	**0.39 (0.24, 0.54)**	118	0.28**	0.93	**0.26 (0.07, 0.42)**	0.13 (-0.10, 0.31)

Note: r= Pearson correlation coefficient; C.b=Bias correction factor; CCC=Bootstrapped Concordance correlation coefficient; CI: Confidence interval; *p-value < 0.05 and **p-value < 0.01

Table 4: Baseline and Week 8 CASI-5 by Child Sex

Characteristics	Boy		Girl		Diff in CCC (Boy - Girl)
	n	CCC (95% CI)	n	CCC (95% CI)	
Baseline CASI-5					
Inattention	80	0.11 (-0.05, 0.25)	28	0.12 (-0.08, 0.34)	-0.01 (-0.27, 0.25)
Hyperactivity	80	**0.33 (0.14, 0.51)**	28	**0.44 (0.13, 0.69)**	-0.11 (-0.43, 0.24)
Impairment	80	**0.46 (0.32, 0.60)**	28	0.28 (-0.13, 0.59)	0.18 (-0.15, 0.60)
ODD	80	**0.44 (0.29, 0.58)**	27	0.21 (-0.12, 0.50)	0.23 (-0.09, 0.58)
DMDD	80	**0.46 (0.30, 0.61)**	27	0.29 (-0.04, 0.56)	0.17 (-0.17, 0.52)
Anger	80	**0.39 (0.18, 0.57)**	27	0.08 (-0.16, 0.45)	0.31 (-0.15, 0.62)
Week 8 CASI-5					
Inattention	73	**0.27 (0.06, 0.46)**	22	**0.60 (0.40, 0.75)**	**-0.33 (-0.58, -0.04)**
Hyperactivity	73	**0.50 (0.34, 0.64)**	22	**0.57 (0.32, 0.78)**	-0.07 (-0.35, 0.20)
Impairment	72	**0.43 (0.22, 0.61)**	22	**0.52 (0.14, 0.81)**	-0.09 (-0.47, 0.34)
ODD	73	**0.45 (0.29, 0.60)**	22	**0.50 (0.12, 0.79)**	-0.05 (-0.39, 0.34)
DMDD	73	**0.24 (0.02, 0.46)**	22	**0.62 (0.19, 0.91)**	-0.38 (-0.77, 0.09)
Anger	73	**0.52 (0.35, 0.66)**	22	0.24 (-0.10, 0.67)	0.28 (-0.16, 0.65)

Note: CCC=Bootstrapped Concordance correlation coefficient; CI: Confidence interval

Table 5: Baseline and Week 8 PROMIS by Child Sex

Characteristics	Boy		Girl		Diff in CCC (Boy - Girl)
	n	CCC (95% CI)	n	CCC (95% CI)	
Baseline PROMIS					
Anger	91	**0.27 (0.09, 0.44)**	34	0.13 (-0.13, 0.37)	0.14 (-0.16, 0.44)
Anxiety	92	0.08 (-0.19, 0.33)	34	0.05 (-0.23, 0.34)	0.03 (-0.35, 0.40)
Depression	92	0.15 (-0.09, 0.41)	34	0.23 (-0.11, 0.54)	-0.08 (-0.47, 0.36)
Peer Relation	89	**0.46 (0.33, 0.61)**	33	0.00 (-0.33, 0.32)	**0.46 (0.08, 0.81)**
Sleep	90	**0.33 (0.12, 0.52)**	33	**0.53 (0.28, 0.72)**	-0.20 (-0.47, 0.12)
Week 8 PROMIS					
Anger	86	**0.24 (0.06, 0.41)**	33	0.21 (-0.03, 0.48)	0.03 (-0.30, 0.32)
Anxiety	87	0.21 (-0.00, 0.42)	33	0.09 (-0.23, 0.39)	0.12 (-0.25, 0.50)
Depression	86	0.17 (-0.08, 0.40)	33	-0.09 (-0.37, 0.24)	0.26 (-0.15, 0.63)
Peer Relation	85	**0.22 (0.03, 0.40)**	33	0.32 (-0.07, 0.67)	-0.10 (-0.54, 0.31)
Sleep	86	**0.36 (0.14, 0.53)**	32	0.07 (-0.35, 0.46)	0.29 (-0.11, 0.77)

Note: CCC=Bootstrapped Concordance correlation coefficient; CI: Confidence interval

Table 6: Summary of Parent-Other Adult Ratings CASI-5 Domains

Characteristics	Baseline				Week 8			
	Parent	Other Adult	Diff	t	Parent	Other Adult	Diff	t
Inattention	2.23 (0.48)	1.79 (0.61)	**0.45**	6.52**	1.78 (0.57)	1.66 (0.61)	0.12	1.80
Hyperactivity	1.75 (0.69)	1.51 (0.88)	**0.24**	2.91**	1.49 (0.70)	1.34 (0.80)	**0.15**	2.03*
Impairment	1.52 (0.64)	1.24 (0.83)	**0.28**	3.73**	1.16 (0.54)	1.09 (0.75)	0.06	0.93
ODD	1.67 (0.69)	1.28 (0.89)	**0.39**	4.68**	1.26 (0.64)	1.06 (0.84)	**0.19**	2.52*
DMDD	1.22 (0.81)	0.92 (0.89)	**0.29**	3.44**	0.85 (0.67)	0.71 (0.80)	0.15	1.70
Anger	0.55 (0.58)	0.53 (0.57)	0.03	0.45	0.35 (0.44)	0.43 (0.51)	-0.08	-1.61

Parent and Other Adult: mean (standard deviation); Diff: average gap between ratings; t: dependent t-test
*p-value < 0.05 and **p-value < 0.01

Table 7: Summary of Parent-Child Ratings PROMIS Domains

Characteristics	Baseline				Week 8			
	Parent	Child	Diff	t	Parent	Child	Diff	t
Anger	2.82 (0.95)	2.38 (0.80)	**0.43**	4.59**	2.46 (0.86)	1.97 (0.70)	**0.50**	5.82**
Anxiety	1.96 (0.70)	2.11 (0.70)	-0.14	-1.68	1.71 (0.62)	1.80 (0.76)	-0.09	-1.11
Depression	1.94 (0.77)	1.90 (0.71)	0.04	0.46	1.67 (0.75)	1.63 (0.59)	0.04	0.53
Peer Relation	3.72 (0.82)	3.65 (0.83)	0.07	0.77	3.97 (0.78)	3.81 (0.98)	0.16	1.57
Sleep	2.68 (0.83)	2.75 (0.71)	-0.07	-0.93	2.35 (0.74)	2.50 (0.76)	**-0.17**	-2.06*

Parent and Child: mean (standard deviation); Diff: average gap between ratings; t: dependent t-test
*p-value < 0.05 and **p-value < 0.01

Table 8: Summary of Parent-Other Adult Ratings CASI-5 Domains by Child Sex

Characteristics	Baseline						Week 8					
	Boy			Girl			Boy			Girl		
	Parent	Other	Diff	Parent	Other	Diff	Parent	Other	Diff	Parent	Other	Diff
Inattention	2.26	1.81	**0.45****	2.16	1.71	**0.45****	1.80	1.70	0.10	1.72	1.52	0.20
Hyperactivity	1.79	1.53	**0.25***	1.66	1.44	0.22	1.47	1.39	0.08	1.54	1.17	**0.37***
Impairment	1.58	1.31	**0.27****	1.36	1.06	0.29	1.16	1.15	0.01	1.13	0.90	0.23
ODD	1.72	1.3	**0.42****	1.52	1.23	0.29	1.28	1.08	**0.20***	1.19	1.01	0.18
DMDD	1.26	0.91	**0.35****	1.06	0.93	0.13	0.86	0.71	0.16	0.82	0.70	0.11
Anger	0.57	0.61	-0.05	0.51	0.27	0.25	0.35	0.50	**-0.15****	0.34	0.20	0.14

Parent and Other Adult: mean ratings; Diff: difference in mean ratings
*p-value < 0.05 and **p-value < 0.01

Table 9: Summary of Parent-Child Ratings PROMIS Domains by Child Sex

Characteristics	Baseline						Week 8					
	Boy			Girl			Boy			Girl		
	Parent	Child	Diff	Parent	Child	Diff	Parent	Child	Diff	Parent	Child	Diff
Anger	2.81	2.37	**0.44****	2.84	2.41	**0.43***	2.47	1.95	**0.53****	2.44	2.00	**0.43***
Anxiety	1.97	2.09	-0.12	1.95	2.16	-0.21	1.70	1.75	-0.05	1.74	1.94	-0.20
Depression	1.96	1.88	0.08	1.88	1.95	-0.07	1.70	1.60	0.09	1.61	1.69	-0.08
Peer Relation	3.73	3.54	**0.20***	3.66	3.95	-0.30	3.94	3.80	0.14	4.06	3.84	0.21
Sleep	2.65	2.74	-0.09	2.75	2.78	-0.02	2.30	2.44	-0.14	2.42	2.68	-0.25

Parent and Child: mean ratings; Diff: difference in mean ratings
*p-value < 0.05 and **p-value < 0.01

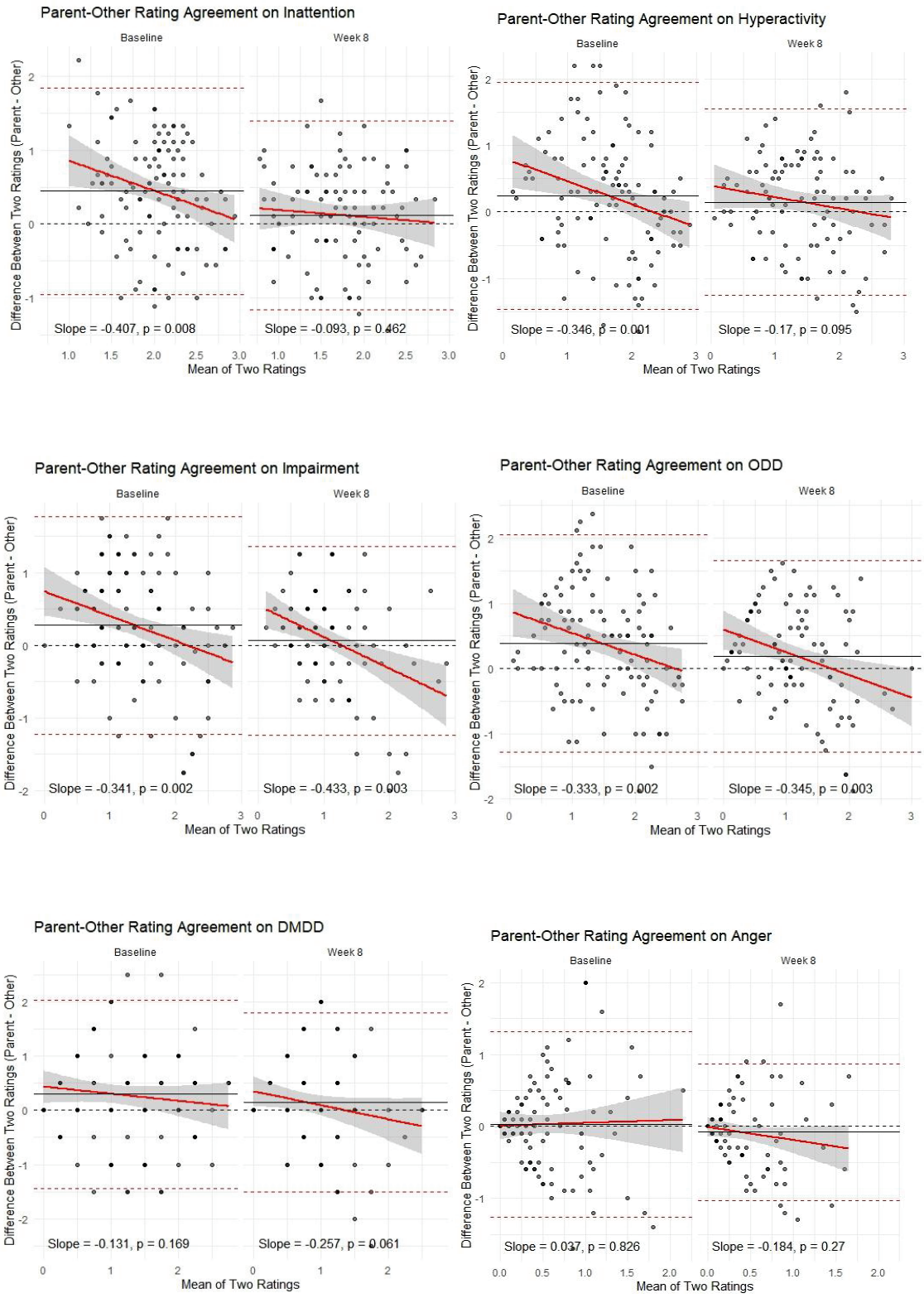

Figure1: Bland-Altman Plots for Parent-Other Adult Agreement Across CASI-5 Domains at Baseline and Week 8

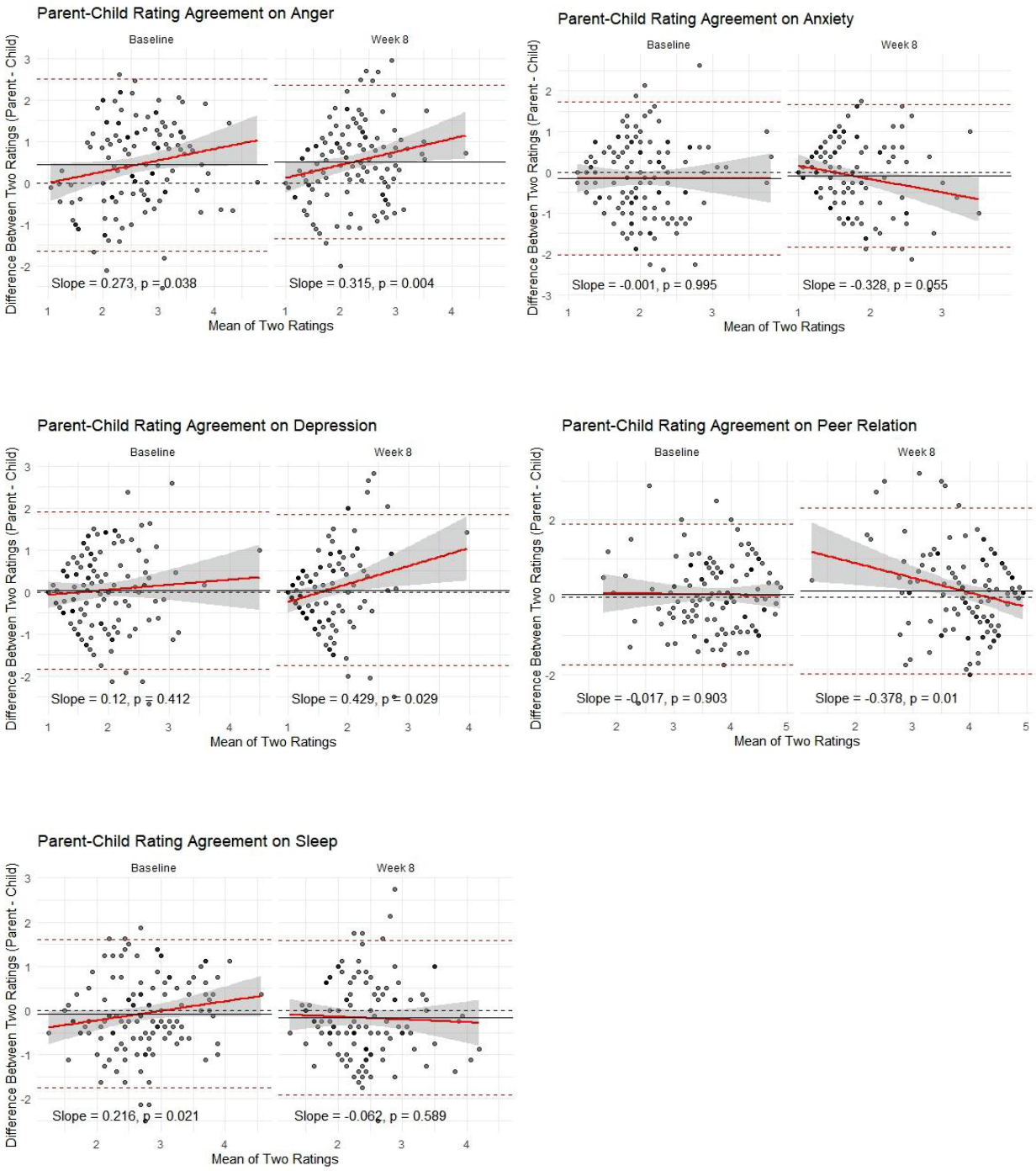

Figure 2: Bland-Altman Plots for Parent-Child Agreement Across PROMIS Domains at Baseline and Week 8

CHAPTER 5: DISCUSSION

In the present study, we examined agreement between parents and another adult (e.g., a teacher, grandparent, aunt/uncle, or neighbor) familiar with the child's behavior, as well as agreement between parents and children's self-ratings, on ADHD core symptoms and co-occurring emotional and behavioral symptoms in a sample of 6-12-year-old children with ADHD and irritability. Parent and other-adult agreement were generally low at both baseline and week 8 for inattention, hyperactivity, social and academic impairment, ODD, DMDD, and anger. However, agreement for inattention significantly improved from baseline to week 8. Parent-child agreement was also low at both time points for anger, anxiety, depression, peer relationships, and sleep issues. Parents tended to report higher symptom severity than other adults across most CASI-5 domains at both time points. Parents also reported significantly greater anger symptoms levels at both baseline and week 8, and significantly lower sleep disturbance at week 8 than their children. Sex-specific differences in agreement were also observed. Girls demonstrated significantly higher parent-other adult agreement than boys for inattention at week 8, whereas boys showed significantly higher parent-child agreement for peer relations at baseline compared with girls.

5.1 Parent-Other Adult Rating Agreement

The CASI-5 measure was used in this study to assess core ADHD symptoms (inattention and hyperactivity), as well as social and academic impairment, ODD, DMDD, and anger issues in children. Both parents and other adults completed CASI-5 measures at baseline and week 8. Across both time points, parent-other adult agreement on CASI-5 domains was low. These findings align with prior research demonstrating low cross-informant agreement for child psychopathology. For example, a meta-analysis of 119 studies by Achenbach et al. (1987)

reported an average parent-teacher correlation of 0.27 on measures of behavioral and emotional problems in children aged 1 to 19 years. By week 8, agreement between parents and other adults on most CASI-5 symptom domains improved, with the significant improved agreement seen in ratings of inattention symptoms. Although overall agreement levels at week 8 across CASI-5 domains remained poor, this positive shift may reflect the influence of the multinutrient intervention, which more than half of participants (56.3%) received. To support the significantly improved agreement on inattention domain, prior research indicated that multinutrient supplementation can reduce and stabilize inattention symptom severity over time (Johnstone et al., 2022; Leung et al., 2024; Rucklidge et al., 2018). Consequently, both parents and other adults likely observed similar behavioral patterns in the children, resulting in notably higher agreement for inattention by week 8. The following details the parent-other adult agreement dynamics across each of the six CASI-5 domains.

5.1.1 Inattention and Hyperactivity

Agreement between parent and other adult ratings of ADHD core symptom domains were poor at both baseline and week 8. These findings are consistent with previous research (Antrop et al., 2002; Mitsis et al., 2000; Murray et al., 2007; Sayal & Goodman, 2009; Sollie et al., 2013; Wolraich et al., 2004), all of which showed that parent-teacher agreement in ADHD symptoms is often low. For example, in a clinically referred sample of children aged 7-11 years, Mitsis et al. (2000) reported low parent-teacher correlations for inattention ($r = 0.30$, $p < .01$) and hyperactivity-impulsivity ($r = 0.39$, $p < .001$). Similarly, in a longitudinal study of elementary school children, Wolraich et al. (2004) found low parent-teacher agreement, with correlations of $r = 0.34$ for inattention and $r = 0.27$ for hyperactivity-impulsivity. Comparable findings have also been observed in a study that included mother, father and teacher rating for agreement

assessment. In that study of children aged 6-15 years referred to a child psychiatric clinic, Sollie et al. (2013) reported significant intraclass correlations between parents and teachers for inattention (mother-teacher ICC = 0.33, $p < .05$; father-teacher ICC = 0.38, $p < .01$) and hyperactivity-impulsivity (mother-teacher ICC = 0.34, $p < .01$; father-teacher ICC = 0.43, $p < .01$). Although the magnitude of parent-other adult agreement for inattention symptoms at baseline was lower than that reported in earlier research, it aligns with a more recent finding, such as Kennerley et al. (2018), who observed a similarly low parent-teacher correlation ($r = 0.14$, $p > .05$). Several factors may account for this reduced level of inattention agreement at baseline. One possibility is the source bias, as parents in our study rated inattentive symptoms as significantly more severe than other adults for both boys and girls (see Table 8), with substantial rating discrepancies (LOA: -0.95 to 1.85) evident at the individual level (see Figure-1). This discrepancy may also partly reflect the less observable and more internalized nature of inattentive symptoms (Gilbert et al., 2025), which may have influenced raters' interpretation of severity. In addition, differences in rater perspectives and expectations about child inattention symptoms may contribute to lower agreement in attention domain (Kennerley et al., 2018). Furthermore, inattentive symptoms are strongly shaped by contextual factors, including task demands, environmental structure, clarity of instructions, and the duration of required attention, all of which may vary substantially between home and outside (e.g., school) settings (Kennerley et al., 2018). The internalizing aspect of inattentive symptoms suggests future research should consider child self-report in the assessment to deepen the understanding of inattention characteristics and to capture child experiences that may not be easily observable to parents or other adults. Moreover, this may necessitate the development and validation of age-appropriate

psychometrically strong self-report tools that enable children to accurately describe their attentional experiences to aid screening accuracy in research.

In this study, parent-other adult agreement was higher for hyperactivity symptoms than for inattentive symptoms at both time points, a pattern that aligns with previous research (Mitsis et al., 2000; Sollie et al., 2013). Taken together with our findings, the literature suggest that differences in informant agreement may be driven by how symptom domains manifest and are perceived across everyday contexts (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005; Duhig et al., 2000; Langberg et al., 2010). That is, hyperactive behaviors are typically outwardly expressive and disruptive in nature and therefore tend to draw attention from adults in various settings (e.g., home and school) (Langberg et al., 2010; Sollie et al., 2013). In contrast, inattentive symptoms often present in more subtle ways, such as difficulty sustaining focus or quietly disengaging during tasks. These behaviors may occur without causing disruption and thus may not be as easily noticed. For instance, a child who frequently leaves their seat or talks excessively is more likely to capture the attention of both parents and other adults compared to a child who seems to be listening but is mentally checked out (Kennerley et al., 2018). This study further indicates that informant agreement may be influenced not only by the presence of symptoms but also by the extent to which those behaviors are disruptive or observable within a given context. For instance, prior research has shown that agreement is generally higher for ratings of observable symptoms compared to less observable or internalized symptoms (De Los Reyes & Kazdin, 2005). In this regard, hyperactive behaviors are more likely to be reported by observers, whereas inattentive symptoms may be interpreted as situational or context-specific and, therefore, less consistently recognized by different observers. Future research should focus on exploring how differently context-specific inattentive symptoms present across settings.

Conversely, Antrop et al. (2002) reported poor and non-significant parent-teacher agreement in a clinically diagnosed sample of children aged 6-12 years with ADHD, with intraclass correlation coefficients of 0.01 for both inattention and hyperactivity. One possible explanation for this is the characteristics of the sample: about half of the children were receiving medication for ADHD symptoms and they often take medication at school (Antrop et al., 2002). As a result, teachers may be more likely to observe the benefits of medication, while parents may not. Additionally, teachers are more familiar with age-appropriate behaviors and may therefore be more tolerant towards problem behaviors than parents (Antrop et al., 2002). Moreover, the structured activities in classroom environments may help children with ADHD regulate their behavior more effectively at school than at home, contributing to lower teacher-reported impairment (Antrop et al., 2002), resulting in low parent-teacher agreement in that study.

Parents tended to rate their children higher than other adult informants on both inattention and hyperactivity symptoms at baseline. Similar trend has been observed at week 8 too. These results are consistent with findings from Antrop et al. (2002), who also reported higher parent ratings compared to teacher ratings for inattention ($t = 4.08$, $p < .01$) and hyperactivity ($t = 3.12$, $p < .01$). Similar findings have been reported by Kennerley et al. (2018) that parents rated significantly more hyperactive/impulsive symptoms ($t = 3.07$, $p < .01$) and total symptoms (inattention and hyperactivity combined) ($t = 2.66$, $p < .01$) than teachers. In addition, Mitsis et al. (2000) and Sollie et al. (2013) also reported higher parents rating than teachers rating. These results are not surprising, as clinical referrals are more likely to occur when parents perceive greater difficulties (Mitsis et al., 2000). Furthermore, higher parental investment in the referral process has been linked to elevated symptom ratings during the assessment (Sayal & Goodman, 2009). Parents are generally deeply committed to their child's long-term outcomes, they may

seek assessments to obtain early interventions, financial support, and services that could promote better future functioning (Mash & Hunsley, 2005; Sonuga-Barke et al., 2011). In contrast, teachers may focus more on classroom behaviors that are disruptive and directly affect classroom management and peer learning. Additionally, teachers, due to their training, are more likely to recognize age-appropriate child behavior and tend to be more tolerant toward disruptive behavior in the classroom setting (Antrop et al., 2002). As a result, teachers may report symptom severity lower relative to parents. Collectively, these may account for the greater discrepancies observed between parent and other adult ratings in the present study. Future research should explore referral mechanisms to better understand why parents report higher symptom levels than other adults for clinically referred sample.

5.1.2 Impairment

In our study, parent-other adult agreement for social and academic impairment was poor at both baseline and week 8. The agreement stayed stable over time. This finding aligns with previous research. For example, McConaughy et al. (2011), reported parent-teacher ratings correlation on social skills with similar magnitude ($r= 0.47$, $p < .001$) in a sample of 6-11 years old children with ADHD. Wolraich et al., (2004) reported low correlation of $r=0.31$ between parent and teacher reports on children academic and classroom performance. Likewise, in a study of children aged 5-12 diagnosed with ADHD, Tamm et al. (2021) found that parent ratings of impairment in academic areas (reading, writing, and math) were significantly and positively correlated with teacher ratings of impairment in the same domains ($r = 0.46–0.69$, $p < .05$). They also observed significant parent-teacher rating correlation for peer relationships ($r=0.31$, $p < .05$). In addition, Sollie et al. (2013) reported similar level of inter-rater agreement for social problems

(mother vs teacher ICC =0.52, p<0.01) and rule breaking behavior (mother vs teacher ICC =0.54, p<0.01).

At baseline, parents tended to rate their child's social and academic impairment as greater than did other adults. By week 8, this discrepancy was no longer evident, which may explain the improvement in parent-other adult agreement over time (see Table 2). Sollie et al. (2013) also reported higher mothers rating on children's social problems and rule breaking behavior than teacher's ratings. To demonstrate this discrepancy, parents may be more likely to notice social difficulties in their child because they observe them across a variety of unstructured and everyday situations such as playing with siblings, interacting with family members, spending time with friends. In these settings, social challenges, such as difficulty making friends, engaging in physical fight, or being rejected by peers, become more visible. Parents are deeply committed to their child's long-term well-being (Mash & Hunsley, 2005) and they perhaps are the first to recognize when their child is struggling socially. In contrast, teachers typically see the child in a structured classroom environment where activities and expectations are clearly defined, and they compare the child only with their classmates (Quinn & Madhoo, 2014). As a result, social difficulties may be less noticeable by the teachers in the structured school setting, leading teachers to rate the child's social impairment as less severe than parents do.

Although teachers are well positioned to evaluate a child's academic performance (Wolraich et al., 2004), their ratings may be influenced primarily by observable subject-related skills. In contrast, parents observe their child completing homework and other learning tasks at home that require independent organization, sustained attention, and self-regulation. These home-based situations may reveal frustration, avoidance, and emotional reactions to academic demands that may not be as apparent in the more structured school environment. Prior research

has also shown that academic difficulties among children with ADHD are often attributable to inattention and reduced productivity (which might be first observed by parents) rather than to deficits in academic skills (McConaughy et al., 2011). Furthermore, parents typically receive more feedback from teachers than teachers receive from parents (Murray et al., 2007). Taken together, these factors may shape parents' perceptions of academic impairment, leading them to report higher levels of difficulty than teachers. In the Casi-5, parents and other adult informants reported how often each symptom domain interfered with their children's social and academic functioning. However, this study did not examine the specific types of social and academic impairments linked to different symptom domains. Future research should therefore assess impairment separately across distinct ADHD domains to better understand how particular symptom domain uniquely affect functioning and whether certain symptom profiles are associated with specific social or academic difficulties.

5.1.3 ODD

In our study, parent-other adult agreement for ODD at baseline was slightly higher than for ADHD core symptoms, consistent with findings reported by Antrop et al. (2002) and Sollie et al. (2013). This higher agreement may reflect the greater concreteness of ODD symptoms; for example, physical aggression or anger can be rated more objectively than the symptoms of ADHD (Antrop et al., 2002). Agreement on ODD in our study were found to be poor at baseline and week 8. This result is consistent with previous research. For example, Antrop et al., (2002) found similar magnitude in parent-teacher agreement for ODD (ICC=0.45, $p<0.01$). Similarly, Sollie et al., (2013) reported ICC value for ODD mother vs teacher is 0.44 ($p<0.01$) and father vs teacher is 0.61 ($p<0.01$). The slightly higher agreement between fathers and teachers may reflect shared tendencies to report fewer behavioral difficulties (Antrop et al., 2002; Sollie et al., 2013).

Fathers have been found to rate their children as exhibiting fewer behavior problems than mothers (Sollie et al., 2013) and teachers may be more likely to view mild disruptive behaviors as developmentally typical (Antrop et al., 2002). A meta-analysis by Petric and Szamosközi (2018) further supports these findings, reporting an overall significant parent-teacher correlation of r = 0.26 ($p < .001$) for ODD across both clinical and non-clinical populations. This relatively smaller effect size may be due to several factors, including the inclusion of both clinical and non-clinical samples as well as the wide age range of participants (1.5 to 17 years) considered in the analysis. Mitchison et al. (2022) also reported poor agreement between parents and teachers for two ODD dimensions such as mood related symptoms and defiant behavior.

Parents tended to rate their child's ODD symptoms higher than other adults at both time points. These findings align with prior research. For example, Antrop et al. (2002) reported significantly higher parent ratings compared to teacher ratings ($t = 3.66$, $p < .01$). Similarly, Sollie et al. (2013) found significant rating discrepancies between mothers and teachers ($t = 2.42$, $p < .05$). This discrepancy may occur due to the difference in how symptoms are perceived across settings or parents and other adults interpreted problematic behavior differently (Mitchison et al., 2022). Research suggests that parents generally tend to report the externalizing behavior more severe than teachers (De Los Reyes et al., 2009; Kerr et al, 2007; Olson et al., 2018). It is also important to consider that informants' knowledge and expectations of typical child development, as well as their relationship with the child, may contribute to rating discrepancies (Mitchison et al., 2022; Petric & Szamosközi, 2018). Future research should examine how environmental context and relationship dynamics influence these differences to better understand the level of parent-other adult agreement in various dimensions of ODD.

5.1.4 DMDD

In our study, parent-other adult agreement for DMDD at both baseline and week 8 were poor. Similar findings have been reported in a study of children aged 6-16 years with Psychiatric Disorders where investigator reported poor parent-teacher agreement on DMDD symptoms (Mayes et al., 2016). In another study examining two aspects of emotional dysregulation (emotional regulation and liability/negativity) among school aged children (5-7 years), investigators reported agreement in similar magnitude between parent and teacher rating (ICC= 0.337 to 0.566) (Mitchison et al., 2022). These findings may indicate that emotional dysregulation is context dependent. The stronger agreement in emotional regulation observed between teacher-teacher reports compared to parent-teacher reports (Mitchison et al., 2022) supports this possibility. Furthermore, differences in subjective ratings and interpretations among informants may contribute to the low level of agreement for DMDD.

Parents tended to rate their child's DMDD symptoms higher than other adults. Mayes et al. (2016) observed similar findings in their study. Mitchison et al. (2022) reported that parents provided significantly higher ratings than teachers on the lability/negativity (e.g., exhibits wide mood swing or recovery from upset or distress) subscale. Parents rate emotional dysregulation higher than teachers or other adults perhaps because it occurs more at home than at school. children tend to mask or supress negative emotional expression in front of peers (Underwood, 1997), leading to fewer observable dysregulation episodes in the school. In addition, parents tend to rate externalizing behavior higher than teachers do (De Los Reyes et al., 2009; Kerr et al., 2007; Olson et al., 2018). As a result, parents may rate emotional dysregulation as more severe than teachers do. In our study, we considered two aspects: whether the child is irritable most of the day and whether temper outbursts are disproportionate to the situation to measure DMDD as

a single construct. Future research should examine distinct aspects of emotional dysregulation in greater depth, investigating how these features manifest across different settings and how these variations influence the level of agreement between parent and other adult.

5.1.5 Anger

In our study, agreement between parents and other adult informants on anger-related behaviors (e.g., annoying others, making threats, provoking physical fights, grabbing, and damaging peers' property) was low at both time points, though it was slightly higher at week 8 compared to baseline. Similar patterns have been reported in previous research. For example, Antrop et al. (2002) found low parent-teacher agreement for disruptive behavior disorders, such as conduct disorder which includes persistent aggression, property destruction, and theft with an ICC of 0.29 ($p < .05$). Likewise, Sollie et al., (2013) reported similar low parent-teacher agreement for aggressive behavior in children with ADHD, with ICC value of 0.30 ($p < .05$) for mothers vs. teachers. To demonstrate the slightly stronger agreement at week 8, one possible explanation might be that since multinutrient intake helps reducing anger related issues and improve overall functioning (Rucklidge et al., 2018), both parents and other adults may observe more consistent anger related behavioral patterns, leading to improved agreement at week 8.

Parents and other adults provided nearly identical mean ratings of anger related behavior at both baseline and week 8. This finding is consistent with previous research indicating minimal informant differences in ratings of overt disruptive behaviors. For example, Sollie et al. (2013) reported no significant mean differences in aggressive behavior ratings between mothers and teachers or between fathers and teachers. Moreover, Antrop et al. (2002) similarly reported no significant parent-teacher rating differences on CD symptoms. One possible explanation is that anger related disruptive behavior is typically showed through overt, outwardly expressed cues

such as verbal outbursts, loud noises, facial expressions, and aggressive gestures (Kerr & Schneider, 2008) that are easily observed across home and school contexts, resulting in lower parent-other adult mean rating differences. In addition, concrete disruptive behaviors (e.g., physical fighting, theft, or property destruction) are more easily and objectively rated, which may further contribute to the similarity in parent and teacher ratings (Antrop et al., 2002). Future research should examine the factors that contribute to variations in anger-related behaviors across different settings and explores strategies to manage and reduce negative impacts on children's social and mental outcomes.

5.2 Parent-Child Rating Agreement

In our study, parent-child rating agreement has also been explored on PROMIS domains (e.g., anger, anxiety, depression, peer relation and sleep). The results indicate poor agreement between self report and parent report. Similar pattern of association (r=0.20) has been reported in a meta-analysis of broadband measures of behavioral and emotional problems in children aged 1 to 19 years (Achenbach et al., 1987). The parent-child agreement across the five PROMIS domains is described below.

5.2.1 Anger

In our study, parent-child agreement on anger symptoms was low at both baseline and week 8. The level of agreement remained stable over the period. Similar pattern of association has been observed in previous research. For example, in a sample of children with diverse acute and chronic health conditions, Varni et al. (2015) reported parent-child correlation (r = 0.40) for anger-related behaviors, reflecting similarly limited agreement. A possible reason for a slightly stronger correlation in that study is that Varni et al. (2015) 's sample included older children

(ages 8–17), who may share a more similar understanding of emotional experiences with parents. In this aspect, past research suggest that older children are typically better able to identify and communicate emotional experiences (Kerr & Schneider, 2008; Riddell et al., 2024), which may lead to higher correlation with parent ratings.

Parents tended to rate children's anger severity higher than children did at both baseline and week 8. This pattern suggests there is perception difference between child and their parent in reporting anger related behavior. It is also possible that parents are more aware of the outward expression and behavioral display of anger of their children than children themselves are. To support this, past research indicates anger is often outwardly expressed through noticeable behaviors such as words or noises, facial expressions, physical gestures, or aggressive movements (Kerr & Schneider, 2008), which may be readily noticed by parents. In contrast, children may underreport anger because they may lack the emotional insight to accurately identify anger or may not perceive their behaviors as problematic (Kerr & Schneider, 2008). However, Varni et al. (2015) reported negligible difference in average symptom score (-0.01), suggesting closer alignment in mean ratings in a sample of children with a diverse health condition. This difference may happen due to the inclusion of older participants in that study. It is important to note that difference between children's and parents' perceptions of the child's symptoms may lead to under-treatment when parents fail to recognize the presence or severity of symptoms, or to over-treatment when parents perceive the symptoms as more severe than the child experiences (Varni et al., 2015). Research indicates emotional experiences are better recognized and communicated by older children (Kerr & Schneider, 2008; Riddell et al., 2024). Future studies should therefore investigate how age influences the expression of anger and

whether developmental differences contribute to informant discrepancies across child age groups.

5.2.2 Anxiety

Prior research indicates that informant agreement tends to be lower for internalizing symptoms such as anxiety or depression than for externalizing behaviors (e.g., aggression, hyperactivity), as the former are generally less observable (Achenbach et al., 1987; Duhig et al, 2000; De Los Reyes & Kazdin, 2005). Consistent with this literature, we found low parent-child agreement on anxiety symptoms at both baseline and week 8. These findings also supported by Kemper et al. (2003), who similarly reported low mother-child agreement on internalizing symptoms, even when both informants completed parallel assessment measures.

In contrast, several studies have reported comparatively stronger parent-child agreement for internalizing symptoms. For instance, in a study of quality-of-life measures Klassen et al. (2006) observed an intraclass correlation of 0.43 for anxiety and depression related mental health items among children with ADHD, and Varni et al. (2015) reported a correlation of $r = 0.41$ for anxiety in a sample of children with diverse acute and chronic health conditions. These higher correlation estimates may be attributable to differences in sample characteristics. Klassen et al. (2006) examined children referred to an ADHD-specialized clinical setting, where many cases involved long-standing, complex diagnostic or management difficulties and prior contact with multiple specialists. In such contexts, parent-child concordance is likely to be higher because both parties have been repeatedly exposed to structured diagnostic evaluations and clinical feedback. Over time, this shared clinical experience may develop increased insight into the child's emotional and behavioral difficulties. Additionally, both Klassen et al. (2006) and Varni et al. (2015) included slightly older participants (ages 10-17 and 8-17 respectively), which may

influence informant agreement, as older children are typically better able to recognize, articulate, and evaluate their internal emotional experiences (Riddell et al., 2024).

On average, children rated their anxiety score slightly higher than parents at both baseline and week 8, although these differences were not statistically significant. This is consistent with findings by Varni et al. (2015), who reported a small difference (0.15) in average symptom score, indicating child reported anxiety score slightly higher than their parents. On the other hand, Klassen et al. (2006) observed that child reported their anxiety and depression related mental health significantly better than parents (mean difference = 6.8, 95% CI: 1.6-12.0, $p < .01$). One possible explanation is that children with ADHD may exhibit positive illusory biases, which lead them to adopt self-protective coping strategies, concealing some problem behaviors to avoid negative parental reactions (Klassen et al., 2006; Owens et al., 2007). Additionally, prior research indicates potential interaction effects between mental health status and comorbid ADHD-related conditions in shaping informant discrepancies. For example, children diagnosed with ODD or CD have been found to rate their mental health and behavior more positively than their parents. Moreover, those presenting with predominantly inattentive or combined ADHD symptom profiles often report higher self-esteem compared to parent ratings (Klassen et al., 2006). These findings suggest that comorbid conditions may systematically influence how children and parents perceive and report psychological functioning. Future research should further investigate how core ADHD symptoms interact with co-occurring internalizing difficulties like anxiety to better understand the underlying informant discrepancies.

5.2.3 Depression

Like anxiety, parent-child agreement on depressive symptoms was also poor at both baseline and week 8. This suggests that parents and children differed in how they perceived or

reported the child's depressive symptoms. Depressive symptoms are often internal and less observable than externalizing behaviors, making them more difficult for parents to detect, resulting in poor informant agreement (Achenbach et al., 1987; Duhig et al, 2000; De Los Reyes & Kazdin, 2005). Moreover, children with ADHD may underreport or have limited insight into their own depressive symptoms (Fraser et al., 2018). In contrast, Varni et al. (2015) observed modest parent-child correspondence (r = 0.37) in a sample of older children (ages 8–17). This older age range may contribute to higher correlation.

Bland-Altman analyses showed that the mean ratings of depressive symptoms provided by parents and children were nearly identical at both baseline and week 8, indicating no systematic bias in either direction. However, both informants reported low depression scores, reflecting the less observable nature of depressive symptoms or the limited insight each informant may have into the child's internal experiences (Achenbach et al., 1987; Duhig et al, 2000; De Los Reyes & Kazdin, 2005; Fraser et al., 2018). Varni et al. (2015) also observed relatively smaller difference (0.13) in average symptoms score for depressive symptoms compared to other PROMIS pediatric domains that are more observable (i.e., tired and pain interference). Future research should explore strategies aimed at enhancing symptom insight and emotional literacy to improve parent-child agreement in the assessment of depressive symptoms.

5.2.4 Peer Relations

In our study, parent-child agreement on peer relations was low at both baseline and week 8. These findings are consistent with prior research. For example, Limbers et al. (2011) reported poor parent-child agreement on social functioning (ICC = 0.29, $p < .001$) among children aged 5-18 years with physician diagnosed ADHD and comorbid psychiatric disorders. Varni et al. (2015) also found similar correlation between parent and child ratings of peer relationships (r = 0.35)

with parent-child mean ratings nearly the same (difference in mean rating=-0.09). Similarly, in the present study, no significant systematic differences were observed between parent and child ratings at either baseline or week 8 (see Table 7). However, despite these comparable mean scores, the low level of agreement may reflect limitations in the accuracy of child self-report. Research indicated that children with ADHD often have considerable difficulty accurately assessing their own social behavior (Hoza et al., 2000), which may contribute to discrepancies between their self-perceptions and parental assessments.

In contrast, in a study of quality-of-life measures among children aged between 8-18 years diagnosed with ADHD, Sciberras et al. (2011) found that child rated their social functioning significantly higher than their parents (mean difference = -16.49, 95% CI: -26.41 to -6.57, $p < .001$). Limbers et al. (2011) also reported similar findings that children (mean age is 11.08 years) rated their social functioning significantly higher than their parents (mean diff=7.82, $p<0.001$). One possible explanation is that, in both studies, the study sample was slightly older than ours. Peer relationships of older children are often less observable to parents (De Los Reyes & Kazdin, 2005) because, as children grow, they spend more time interacting with peers outside the home such as at school, playground or extracurricular activities. As a result, parents may have less accurate insight into their child's peer relationships and related challenges, contributing to higher discrepancies between parent and child reports on social functioning and peer-related issues. Given that children with ADHD experience complex and multifaceted peer difficulties (Binti et al., 2022), and that their self-reports of peer relationships often show substantial inconsistencies with observed actual performance (Hoza et al., 2000), future research should consider incorporating peer informants to obtain a more comprehensive understanding of their

social functioning. Evidence suggests that peer evaluations are stronger predictors of later psychiatric difficulties than adult reports (Cowen et al., 1973).

5.2.5 Sleep

In our study, parent-child agreement on sleep disturbance was low at both baseline and week 8. This result aligns with findings in other studies comparing subjective and objective sleep issues of school-aged children (Owens et al., 2009; Owens et al., 2000b). This low agreement likely reflects differences in perception and reporting between informants. Children may overestimate sleep problems due to limited understanding of their sleep patterns and difficulty accurately answering items require time-estimation (Owens et al., 2000b). Conversely, parents may lack full awareness of their child's nighttime experiences or may not interpret certain sleep behaviors as problematic (Owens et al., 2000b). In contrast, Owens et al. (2000a) reported a slightly higher level of correlation (Spearman r = 0.55) between parent and child reported sleep disturbances in a clinical sample of school-aged children (5-10 years) with ADHD. The authors indicated that this elevated correlation might reflect reporter bias, as parents may have assisted their children in completing the survey.

The difference in mean ratings at baseline was not statistically significant. However, by week 8, children on average reported more sleep problems than their parents, which reflects the relatively weaker agreement in week 8 relative to baseline (see Table 3). This pattern aligns with findings by Owens et al. (2000a), who observed that elementary school-aged children tended to report more sleep difficulties particularly sleep-onset delays and night wakings than their parents. This discrepancy may arise because some children do not alert their parents when they wake at night, while others handle trouble falling asleep on their own (such as by reading or watching TV), so parents may not recognize the problem (Owens et al., 2000a; Owens et al., 2000b).

Additionally, over the eight-week period, it is possible that children's greater self-awareness might have led them to recognize and report ongoing or more subtle sleep difficulties that parents did not notice, resulting in a higher discrepancy between parent and child reports at week 8. In this study, we treated sleep disturbance as a single uniform construct. Future research should examine multiple dimensions of sleep disturbances in children with ADHD and assess parent-child agreement within each area to better understand where and why their ratings align or differ.

5.3 Agreement By Child Sex

In the present study, we also explored the level of informant agreement stratified by child sex (see Table 4 and table 5). Consistent with the literature, the pattern of informant agreement differed notably between boys and girls, reflecting well-established gender differences in the presentation of ADHD and related behaviors (Gershon & Gershon, 2002; Gilbert et al., 2025; Loyer Carbonneau et al., 2021). Below we have discussed the difference in informant agreement by child sex for both CASI-5 and PROMIS symptom domains.

5.3.1 CASI-5 Domains

At baseline, boys showed low parent-other adult agreement across all six CASI-5 domains with relative higher agreement in more observable domains (like ODD and DMDD) than less observable domains (like inattention). Boys tend to exhibit more externalizing behaviors (Gershon & Gershon, 2002; Rucklidge 2010), which are more overt and therefore more easily observed across settings (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005). This likely contributes to the higher agreement across these domains. On the other hand, for girls, highest agreement emerged only for hyperactivity perhaps due to its disruptive nature. Consistent with the meta-analysis by Achenbach et al. (1987), the overall difference in agreement

between boys and girls at baseline was not statistically significant in our study. One possible explanation for the slightly higher parent-other adult agreement on hyperactivity among girls compared to boys is the lower base rate of hyperactive behavior in girls (Gershon & Gershon, 2002; Quinn & Madhoo, 2014). When hyperactivity does occur in girls, it may be atypical and more likely to stand out across settings, increasing the likelihood that both parents and other adults notice and report similar levels of symptoms.

At the 8-week assessment, overall symptom severity decreased for both boys and girls (see Table 8), potentially reflecting the positive effects of the multinutrient intervention (Johnstone et al., 2022; Leung et al., 2024; Rucklidge et al., 2014; Rucklidge et al., 2018). Although parent-other adult agreement strengthened across groups at week 8 compared to baseline, girls showed a notable increase in agreement relative to boys, specifically for inattention. This may be explained by the greater prominence of inattentive symptom presentations among girls (Ayano et al., 2023; Gershon & Gershon, 2002; Quinn & Madhoo, 2014). As inattentive symptoms improved following multinutrient supplementation, these changes may have become more readily detectable across both home and school environments, reducing situational variability in symptom expression. Consequently, symptom reductions in girls may have produced more convergent behavioral cues for parents and other adults, thereby strengthening parent-other adult agreement over time. This marked improvement in parent-other adult agreement for inattention among girls at week 8 may therefore have contributed substantially to the stronger agreement observed at week 8 in the overall sample.

The parent-other adult rating discrepancy is more prominent among boys than girls and both parents and other adults consistently rated girls as having lower symptom severity than boys (see Table 8). It may be because females with ADHD may develop stronger coping mechanisms

and are often better able to mask the effects of their symptoms compared to males (Quinn & Madhoo, 2014). As a result, their symptoms among girls may appear less impaired across settings (Gershon & Gershon, 2002), leading to smaller discrepancies between parent and other adult reports. There is limited research examining the factors that contribute to variations in rating discrepancies based on children's sex. Future studies should investigate these factors to enhance understanding of how child sex may shape informant discrepancies.

5.3.2 PROMIS Domains

Across both time points, boys exhibited higher parent-child agreement on anger, peer relations, and sleep relative to internalized symptoms like anxiety and depression. On the other hand, girls showed strongest parent-child agreement in sleep at baseline. This pattern may reflect gender differences in symptom expression. Anger in boys with ADHD tends to be more externalized, often appearing as physical or verbal outbursts (Isaksson et al., 2025) that are readily observable to parents. Girls with ADHD, however, are more likely experiencing higher rates of internalizing symptoms such as depression and anxiety (Gershon & Gershon, 2002; Isaksson et al., 2025), which are closely linked with sleep problems (Dimakos et al., 2024). Consequently, sleep disturbances may be more noticeable and consistently reported for girls at baseline, resulting in higher parent-child agreement in this domain.

Boys in our sample showed significantly higher parent-child agreement on peer relations than girls, although the wide confidence interval suggests considerable uncertainty in the magnitude of this difference. One possible explanation is that peer difficulties among boys may be attributable to overt and disruptive behaviors such as hyperactivity and proactive aggression (Dimitri et al., 2025) that are easily observable by parents. In contrast, girls with ADHD tend to experience more subtle and relational forms of peer impairment, including social exclusion,

relational aggression, and internalized worry about friendships (kok et al., 2016), which may be less visible and therefore less likely to be accurately detected by parents. Additionally, girls have better coping skills than boys (Quinn & Madhoo, 2014) which may mask their social challenges, further reducing parents' awareness of these difficulties. Together, these differences in the visibility and expression of peer problems may explain the lower parent-child agreement observed among girls. However, future studies are needed to examine the consistency of these findings and to further investigate how a child's sex may influence parent-child agreement regarding peer relationships.

Across both time points, parents consistently reported higher levels of anger symptoms in their children than the children reported themselves, for both boys and girls (see Table 9). Because anger is often expressed through observable behaviors such as vocal outbursts, facial expressions, physical gestures, or aggressive actions (Kerr & Schneider, 2008) which may be more readily noticed by parents. At the same time, children may underreport anger due to their limited awareness of internal emotional states (Kerr & Schneider, 2008). At baseline, among boys, parents rated peer-relationship significantly higher than boys' self-reports. This discrepancy may reflect the nature of boys' social interactions, which often occur outside the home such as at school, during extracurricular activities, or on playgrounds where parents have limited opportunity to observe peer dynamics (National Research Council, 1984). Without direct supervision, parents may be less aware of their child's friendships, conflicts, or social challenges. Consequently, it may contribute to informant discrepancies in ratings of social functioning and peer-related difficulties. To better understand the role of child sex in rater agreement, future research should investigate the factors underlying variation in parent-child agreement between boys and girls across PROMIS domains.

5.4 Strength and Limitations

There are several strengths associated with this study. Firstly, recruiting participants from multiple sites reduces the likelihood that findings are driven by site-specific characteristics (e.g., local clinical practices, demographic profiles, referral patterns). This enhances the external validity of the study by ensuring that the sample better represents the diversity of the broader population of children being assessed for ADHD. Secondly, a randomized controlled trial itself a rigorous methodological framework that minimizes selection bias and confounding. Unlike cross-sectional study, collecting rater assessments at multiple time points within an RCT allows for evaluation of stability and change in inter-rater agreement over time. Thirdly, using validated and widely adopted assessment instruments (CASI-5 and PROMIS) ensures reducing measurement error. Additionally, ADHD often co-occurs with emotional and behavioral difficulties, which can influence symptom perception across raters. By including emotional and behavior measures, the study goes beyond core ADHD symptoms and assesses broader psychosocial functioning. This comprehensive approach allows for understanding how agreement varies across symptom domains and supports a more holistic view of children's ADHD and related disorder. Furthermore, stratifying analyses by child sex allows the study to explore whether agreement patterns differ between boys and girls. This approach helps uncover nuanced patterns that might be masked in aggregate analyses. Such findings can contribute to more tailored assessment and screening practices. Finally, although most earlier studies used the Pearson correlation coefficient, with a few incorporating the intraclass correlation coefficient, our analyses examined informant agreement through two different approaches. First, we utilized bootstrapped concordance correlation coefficient (CCC) values to understand the strength of the agreement. Second, we used Bland-Altman plots to explore the existence of systematic biases

and the level of agreement at the individual level. Unlike the Pearson correlation coefficient, which measures only the strength of the linear association between two variables, the CCC combines the Pearson correlation (which reflects how closely the data cluster around a straight line) with a bias correction factor that quantifies how far this line deviates from the line of perfect agreement (the $y = x$ line). In addition, bootstrapping improves estimation accuracy and confidence interval stability. Moreover, Bland-Altman plots offer intuitive visualization of systematic biases and the magnitude of disagreement at the individual level. Together, these methods provide a more comprehensive and statistically robust evaluation of rater agreement across symptom domains than correlation alone.

A limitation of this study is that participants' eligibility was determined based on parent completed CASI-5 questionnaire rather than a diagnostic clinical interview. The study relied primarily on subjective rating scales completed by parents, children, and other adults who knew the child well to assess ADHD and related symptom domains. Subjective reports are vulnerable to rater bias, mood effects, recall variability, and differences in rater knowledge and expectations, which may introduce measurement error and decrease the internal validity of the findings (Szklo & Nieto, 2014). Informants were required to retrospectively recall the child's symptoms and behaviors, which may be influenced by memory decay or current emotional state. Retrospective recall introduces recall bias (Szklo & Nieto, 2014), which can reduce the accuracy of symptom ratings and compromise the reliability of baseline or follow-up estimates. In the present study, parents served as one of primary informants. However, using the gender-neutral term "parents" may obscure meaningful differences in how mothers and fathers perceive and report child difficulties (Sollie et al., 2013). Future research should consider obtaining ratings from both parents to better capture potential differences in perspectives. Performing sex-stratified analyses

with smaller subgroups further reduces statistical power and may produce less precise estimates. Finally, multiple comparisons were not adjusted for, increasing the risk of Type I error; however, these analyses were exploratory and intended to characterize agreement patterns rather than test a single confirmatory hypothesis, therefore, formal multiple-comparison correction was not applied.

5.5 Conclusion

The present study demonstrated that in a sample of 6-12-year-old children with ADHD and irritability, parent-other adult and parent-child agreement on ADHD core symptoms and co-occurring emotional and behavioral problems is generally low, with the strength of agreement varying across symptom domains and differing by child sex. Although our study has several limitations, these findings contribute meaningfully to the existing evidence base. Given there is no single gold-standard measure for assessing psychopathology in children, multi-informant assessment remains essential (De Los Reyes & Kazdin, 2005; Eom & Kim, 2023; Kennerley et al., 2018). However, incorporating multiple informants presents a consistent challenge: rating discrepancies, and such disagreement can negatively affect assessment accuracy, case conceptualization, and subsequent treatment planning for childhood psychopathology (De Los Reyes & Kazdin, 2005). Despite these challenges, informant discrepancies are not inherently problematic; in fact, they can provide valuable information (Achenbach et al., 1987). A comprehensive evaluation that integrates input from multiple informants remains the optimal approach (Mitsis et al., 2000; Wolraich et al., 2004). Discrepant reports can shed light on how a child's behavior varies across settings, offering insight that may be essential for individualized case formulation and treatment planning (Langberg et al., 2010; Mitsis et al., 2000; Sollie et al., 2013). It also reflects variations in contextual demands, informants' knowledge and expectations,

the nature of their relationships with the child, and subjective interpretations of behavioral and emotional symptoms (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005; Duhig et al., 2000; Kennerley et al., 2018; Petric & Szamosközi, 2018; Sollie et al., 2013). This study highlighted the ongoing need for researchers to recognize and assess informant agreement and rating discrepancies and minimize the negative impact this can have on a child's holistic development over time.

REFERENCES

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. https://doi.org/10.1037/0033-2909.101.2.213

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Antrop, I., Roeyers, H., Oosterlaan, J., & Van Oost, P. (2002). Agreement between parent and teacher ratings of disruptive behavior disorders in children with clinically diagnosed ADHD. *Journal of Psychopathology and Behavioral Assessment*, *24*, 67-73. https://doi.org/10.1023/a:1014057325752

Arnold, L. E., Hodgkins, P., Kahle, J., Madhoo, M., & Kewley, G. (2020). Long-term outcomes of ADHD: academic achievement and performance. *Journal of attention disorders*, *24*(1), 73-85. https://doi.org/10.1177/1087054714566076

Ayano, G., Demelash, S., Gizachew, Y., Tsegay, L., & Alati, R. (2023). The global prevalence of attention deficit hyperactivity disorder in children and adolescents: An umbrella review of meta-analyses. *Journal of affective disorders*, *339*, 860-866. https://doi.org/10.1016/j.jad.2023.07.071

Azeredo, A., Moreira, D., & Barbosa, F. (2018). ADHD, CD, and ODD: Systematic review of genetic and environmental risk factors. *Research in developmental disabilities*, *82*, 10-19. https://doi.org/10.1016/j.ridd.2017.12.010

Bingöl-Kızıltunç, P., Yürümez, E., & Atilla, H. (2022). Does methylphenidate treatment affect functional and structural ocular parameters in patients with attention deficit hyperactivity disorder?-A prospective, one year follow-up study. *Indian journal of ophthalmology*, *70*(5), 1664-1668. https://doi.org/10.4103/ijo.ijo_2966_21

Binti Marsus, N., Huey, L. S., Saffari, N., & Motevalli, S. (2022). Peer relationship difficulties among children with ADHD: a systematic review. *International Journal of Academic Research in Business and Social Sciences*, *12*(6), 1265-1276. https://doi.org/10.6007/ijarbss/v12-i6/13352

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical methods in medical research*, *8*(2), 135-160. https://doi.org/10.1191/096228099673819272

Bryant, A., Schlesinger, H., Sideri, A., Holmes, J., Buitelaar, J., & Meiser-Stedman, R. (2023). A meta-analytic review of the impact of ADHD medications on anxiety and depression in children and adolescents. *European Child & Adolescent Psychiatry*, *32*(10), 1885-1898. https://doi.org/10.1007/s00787-022-02004-8

Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., ... & PROMIS Cooperative Group. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item

banks: 2005–2008. *Journal of clinical epidemiology*, *63*(11), 1179-1194. https://doi.org/10.1016/j.jclinepi.2010.04.011

Claussen, A. H., Holbrook, J. R., Hutchins, H. J., Robinson, L. R., Bloomfield, J., Meng, L., ... & Kaminski, J. W. (2024). All in the family? A systematic review and meta-analysis of parenting and family environment as risk factors for attention-deficit/hyperactivity disorder (ADHD) in children. *Prevention Science*, *25*(Suppl 2), 249-271. https://doi.org/10.1007/s11121-022-01358-4

Chan, E., Zhan, C., & Homer, C. J. (2002). Health care use and costs for children with attention-deficit/hyperactivity disorder: national estimates from the medical expenditure panel survey. *Archives of pediatrics & adolescent medicine*, *156*(5), 504-511. https://doi.org/10.1001/archpedi.156.5.504

Cohen-Zion, M., & Ancoli-Israel, S. (2004). Sleep in children with attention-deficit hyperactivity disorder (ADHD): a review of naturalistic and stimulant intervention studies. *Sleep medicine reviews*, *8*(5), 379-402. https://doi.org/10.1016/j.smrv.2004.06.002

Cowen, E. L., Pederson, A., Babigian, H., Isso, L. D., & Trost, M. A. (1973). Long-term follow-up of early detected vulnerable children. *Journal of consulting and clinical psychology*, *41*(3), 438. https://doi.org/10.1037/h0035373

Dalsgaard, S., Kvist, A. P., Leckman, J. F., Nielsen, H. S., & Simonsen, M. (2014). Cardiovascular safety of stimulants in children with attention-deficit/hyperactivity disorder: a nationwide prospective cohort study. *Journal of child and adolescent psychopharmacology*, *24*(6), 302-310. https://doi.org/10.1089/cap.2014.0020

De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of abnormal child psychology*, *37*(5), 637-652. https://doi.org/10.1007/s10802-009-9307-3

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological bulletin, 131*(4), 483. https://doi.org/10.1037/0033-2909.131.4.483

Dickstein, D. P. (2016). Mechanisms distinguishing irritability in children and adolescents. *American Journal of Psychiatry*, *173*(7), 653-654. https://doi.org/10.1176/appi.ajp.2016.16040429

Dimakos, J., Gauthier-Gagné, G., Lin, L., Scholes, S., & Gruber, R. (2024). The associations between sleep and externalizing and internalizing problems in children and adolescents with attention-deficit/hyperactivity disorder: empirical findings, clinical implications, and future research directions. *Psychiatric Clinics*, *47*(1), 179-197. https://doi.org/10.1016/j.psc.2023.06.012

Dimitri, D., Delia, G., Cavallo, F., Varini, M., & Fioretto, F. (2025). Sex differences in children and adolescents with attention-deficit/hyperactivity disorder: a literature review. *Frontiers in Child and Adolescent Psychiatry*, *4*, 1582502. https://doi.org/10.3389/frcha.2025.1582502

Doshi, J. A., Hodgkins, P., Kahle, J., Sikirica, V., Cangelosi, M. J., Setyawan, J., ... & Neumann, P. J. (2012). Economic impact of childhood and adult attention-deficit/hyperactivity

disorder in the United States. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(10), 990-1002. https://doi.org/10.1016/j.jaac.2012.07.008

Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice, 7*(4), 435-453. https://doi.org/10.1093/clipsy.7.4.435

DuPaul, G. J., Volpe, R. J., Jitendra, A. K., Lutz, J. G., Lorah, K. S., & Gruber, R. (2004). Elementary school students with AD/HD: Predictors of academic achievement. *Journal of School Psychology*, *42*(4), 285-301. https://doi.org/10.1016/j.jsp.2004.05.001

Eom, T. H., & Kim, Y. H. (2023). Clinical practice guidelines for attention-deficit/hyperactivity disorder: recent updates. *Clinical and Experimental Pediatrics*, *67*(1), 26. https://doi.org/10.3345/cep.2021.01466

Faraone, S. V., Biederman, J., Morley, C. P., & Spencer, T. J. (2008). Effect of stimulants on height and weight: a review of the literature. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(9), 994-1009. https://doi.org/10.1097/chi.0b013e31817e0ea7

Faraone, S. V., Biederman, J., Weber, W., & Russell, R. L. (1998). Psychiatric, neuropsychological, and psychosocial features of DSM-IV subtypes of attention-deficit/hyperactivity disorder: results from a clinically referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *37*(2), 185-193. https://doi.org/10.1097/00004583-199802000-00011

Forrest, C. B., Meltzer, L. J., Marcus, C. L., De La Motte, A., Kratchman, A., Buysse, D. J., ... & Bevans, K. B. (2018). Development and validation of the PROMIS Pediatric Sleep

Disturbance and Sleep-Related Impairment item banks. *Sleep*, *41*(6), zsy054. https://doi.org/10.1093/sleep/zsy054

Fraser, A., Cooper, M., Agha, S. S., Collishaw, S., Rice, F., Thapar, A., & Eyre, O. (2018). The presentation of depression symptoms in attention-deficit/hyperactivity disorder: comparing child and parent reports. *Child and Adolescent Mental Health*, *23*(3), 243-250. https://doi.org/10.1111/camh.12253

Gadow, K. D., & Sprafkin, J. (2005). *Child and adolescent symptom inventory-4R. Stony Brook, NY: Checkmate Plus*. https://www.checkmateplus.com/

Gadow, K. D., & Sprafkin, J. (2013). *Child and adolescent symptom inventory-5. Checkmate plus. Stony brook*. https://www.checkmateplus.com/

Gadow, K. D., & Sprafkin, J. (2015). The symptom inventories: An annotated bibliography. *Stony Brook, NY: Checkmate Plus*. https://www.checkmateplus.com/

Gadow, K. D., Sprafkin, J., Salisbury, H., Schneider, J., & Loney, J. (2004). Further Validity Evidence for the Teacher Version of the Child Symptom lnventory-4. *School Psychology Quarterly*, *19*(1), 50. https://doi.org/10.1521/scpq.19.1.50.29408

Gershon, J., & 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

Gilbert, M., Boecker, M., Reiss, F., Kaman, A., Erhart, M., Schlack, R., ... & Ravens-Sieberer, U. (2025). Gender and age differences in ADHD symptoms and co-occurring depression and

anxiety symptoms among children and adolescents in the BELLA study. *Child Psychiatry & Human Development*, *56*(4), 1162-1172. https://doi.org/10.1007/s10578-023-01622-w

Gnanavel, S., Sharma, P., Kaushal, P., & Hussain, S. (2019). Attention deficit hyperactivity disorder and comorbidity: A review of literature. *World journal of clinical cases*, *7*(17), 2420. https://doi.org/10.12998/wjcc.v7.i17.2420

Gordon, H. A., Rucklidge, J. J., Blampied, N. M., & Johnstone, J. M. (2015). Clinically significant symptom reduction in children with attention-deficit/hyperactivity disorder treated with micronutrients: an open-label reversal design study. *Journal of child and adolescent psychopharmacology*, *25*(10), 783-798. https://doi.org/10.1089/cap.2015.0105

Graham, J., Banaschewski, T., Buitelaar, J., Coghill, D., Danckaerts, M., Dittmann, R. W., ... & European Guidelines Group). (2011). European guidelines on managing adverse effects of medication for ADHD. *European child & adolescent psychiatry*, *20*, 17-37. https://doi.org/10.1007/s00787-010-0140-6

Graham, J., & Coghill, D. (2008). Adverse effects of pharmacotherapies for attention-deficit hyperactivity disorder: epidemiology, prevention and management. *CNS drugs*, *22*, 213-237. https://doi.org/10.2165/00023210-200822030-00003

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—the gold standard for effectiveness research. *BJOG: an international journal of obstetrics and gynaecology*, *125*(13), 1716. https://doi.org/10.1111/1471-0528.15199

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow

process for providing translational research informatics support. *Journal of biomedical informatics*, *42*(2), 377-381. https://doi.org/10.1016/j.jbi.2008.08.010

Hennissen, L., Bakker, M. J., Banaschewski, T., Carucci, S., Coghill, D., Danckaerts, M., ... & Adduce Consortium. (2017). Cardiovascular effects of stimulant and non-stimulant medication for children and adolescents with ADHD: a systematic review and meta-analysis of trials of methylphenidate, amphetamines and atomoxetine. CNS drugs, 31, 199-215. https://doi.org/10.1007/s40263-017-0410-7

Hoza, B., Waschbusch, D. A., Pelham, W. E., Molina, B. S., & Milich, R. (2000). Attention-deficit/hyperactivity disordered and control boys' responses to social success and failure. *Child development*, *71*(2), 432-446. https://doi.org/10.1111/1467-8624.00155

Huhdanpää, H., Morales-Munoz, I., Aronen, E. T., Pölkki, P., Saarenpää-Heikkilä, O., Kylliäinen, A., & Paavonen, E. J. (2021). Prenatal and postnatal predictive factors for children's inattentive and hyperactive symptoms at 5 years of age: the role of early family-related factors. *Child Psychiatry & Human Development*, *52*, 783-799. https://doi.org/10.1007/s10578-020-01057-7

Isaksson, J., Sukhodolsky, D. G., Koposov, R., Stickley, A., Ramklint, M., & Ruchkin, V. (2025). Attention-Deficit/Hyperactivity Disorder Symptoms and Anger and Aggression in Russian Adolescents. *JAACAP open*, *3*(1), 126-136. https://doi.org/10.1016/j.jaacop.2024.01.006

Johnstone, J. M., Hatsu, I., Tost, G., Srikanth, P., Eiterman, L. P., Bruton, A. M., ... & Arnold, L. E. (2022). Micronutrients for attention-deficit/hyperactivity disorder in youths: a placebo-

controlled randomized clinical trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *61*(5), 647-661. https://doi.org/10.1016/j.jaac.2021.07.005

Johnstone, J. M., Leung, B., Gracious, B., Perez, L., Tost, G., Savoy, A., ... & Arnold, L. E. (2019). Rationale and design of an international randomized placebo-controlled trial of a 36-ingredient micronutrient supplement for children with ADHD and irritable mood: The Micronutrients for ADHD in Youth (MADDY) study. *Contemporary clinical trials communications*, *16*, 100478. https://doi.org/10.1016/j.conctc.2019.100478

Kamal, M., Eltohami, M., Al-Shibli, S., & Khan, S. (2018). Effect and side effect of stimulant/methylphenidate on children and adolescents with ADHD: the Qatar experience. *Arch Pediatr: JPED-160*. DOI: 10.29011/2575-825X.100060

Kaplan, B. J., Steiger, R. A., Pope, J., Marsh, A., Sharp, M., & Crawford, S. G. (2010). Successful treatment of pill-swallowing difficulties with head posture practice. *Paediatrics & child health*, *15*(5), e1-e5. https://doi.org/10.1093/pch/15.5.e1

Kemper, T. S., Gerhardstein, R., Repper, K. K., & Kistner, J. A. (2003). Mother–child agreement on reports of internalizing symptoms among children referred for evaluation of ADHD. *Journal of Psychopathology and Behavioral Assessment*, *25*, 239-250. https://doi.org/10.1023/a:1025847012201

Kennerley, S., Jaquiery, B., Hatch, B., Healey, M., Wheeler, B. J., & Healey, D. (2018). Informant discrepancies in the assessment of attention-deficit/hyperactivity disorder. *Journal of Psychoeducational Assessment*, *36*(2), 136-147. https://doi.org/10.1177/0734282916670797

Kerr, D. C., Lunkenheimer, E. S., & Olson, S. L. (2007). Assessment of child problem behaviors by multiple informants: A longitudinal study from preschool to school entry. *Journal of Child Psychology and psychiatry*, *48*(10), 967-975. https://doi.org/10.1111/j.1469-7610.2007.01776.x

Kerr, M. A., & Schneider, B. H. (2008). Anger expression in children and adolescents: A review of the empirical literature. *Clinical psychology review*, *28*(4), 559-577. https://doi.org/10.1016/j.cpr.2007.08.001

Kim, J. H., Kim, J. Y., Lee, J., Jeong, G. H., Lee, E., Lee, S., ... & Fusar-Poli, P. (2020). Environmental risk factors, protective factors, and peripheral biomarkers for ADHD: an umbrella review. *The Lancet Psychiatry*, *7*(11), 955-970. https://doi.org/10.1016/s2215-0366(20)30312-6

Kitchens, S. A., Rosén, L. A., & Braaten, E. B. (1999). Differences in anger, aggression, depression, and anxiety between ADHD and non-ADHD children. *Journal of Attention Disorders*, *3*(2), 77-83. https://doi.org/10.1177/108705479900300201

Klassen, A. F., Miller, A., & Fine, S. (2006). Agreement between parent and child report of quality of life in children with attention-deficit/hyperactivity disorder. *Child: care, health and development*, *32*(4), 397-406. https://doi.org/10.1111/j.1365-2214.2006.00609.x

Kok, F. M., Groen, Y., Fuermaier, A. B., & Tucha, O. (2016). Problematic peer functioning in girls with ADHD: A systematic literature review. *PloS one*, *11*(11), e0165119. https://doi.org/10.1371/journal.pone.0165119

Langberg, J. M., Epstein, J. N., Simon, J. O., Loren, R. E., Arnold, L. E., Hechtman, L., ... & Wigal, T. (2010). Parental agreement on ADHD symptom-specific and broadband

externalizing ratings of child behavior. *Journal of Emotional and Behavioral Disorders*, *18*(1), 41. https://doi.org/10.1177/1063426608330792

Lenzi, F., Cortese, S., Harris, J., & Masi, G. (2018). Pharmacotherapy of emotional dysregulation in adults with ADHD: a systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *84*, 359-367. https://doi.org/10.1016/j.neubiorev.2017.08.010

Leung, B. M., Srikanth, P., Robinette, L., Bruton, A. M., Tost, G., Hatsu, I., ... & Johnstone, J. M. (2024). Micronutrients for ADHD in youth (MADDY) study: comparison of results from RCT and open label extension. *European Child & Adolescent Psychiatry*, *33*(5), 1355-1367. https://doi.org/10.1007/s00787-023-02236-2

Limbers, C. A., Ripperger-Suhler, J., Heffer, R. W., & Varni, J. W. (2011). Patient-reported Pediatric Quality of Life Inventory™ 4.0 Generic Core Scales in pediatric patients with attention-deficit/hyperactivity disorder and comorbid psychiatric disorders: feasibility, reliability, and validity. *Value in Health*, *14*(4), 521-530. https://doi.org/10.1016/j.jval.2010.10.031

Lin, L., Hedayat, A. S., Sinha, B., & Yang, M. (2002). Statistical methods in assessing agreement: Models, issues, and tools. *Journal of the American Statistical Association*, *97*(457), 257-270. https://doi.org/10.1198/016214502753479392

Loyer Carbonneau, M., Demers, M., Bigras, M., & Guay, M. C. (2021). Meta-analysis of sex differences in ADHD symptoms and associated cognitive deficits. *Journal of attention disorders*, *25*(12), 1640-1656. https://doi.org/10.1177/1087054720923736

Lycett, K., Mensah, F. K., Hiscock, H., & Sciberras, E. (2014). A prospective study of sleep problems in children with ADHD. *Sleep Medicine*, *15*(11), 1354-1361. https://doi.org/10.1016/j.sleep.2014.06.004

Martel, M. M., Schimmack, U., Nikolas, M., & Nigg, J. T. (2015). Integration of symptom ratings from multiple informants in ADHD diagnosis: a psychometric model with clinical utility. *Psychological assessment*, *27*(3), 1060. https://doi.org/10.1037/pas0000088

Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of clinical child and adolescent psychology*, *34*(3), 362-379. https://doi.org/10.1207/s15374424jccp3403_1

Mayes, S., Waxmonsky, J., Waschbusch, D., Mattison, R., Baweja, R., Hameed, U., & Syed, E. (2016). Mother, father, and teacher agreement on disruptive mood dysregulation disorder symptoms in children with psychiatric disorders. *International Journal of Mental Health & Psychiatry*, *2*(2). https://doi.org/10.4172/2471-4372.1000123

McBride, G. B. (2005). A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. *NIWA client report: HAM2005-062*, *45*, 307-310. https://www.medcalc.org/download/pdf/McBride2005.pdf.

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. https://doi.org/10.1080/02796015.2011.12087713

Mitchison, G. M., Liber, J. M., & Njardvik, U. (2022). Parent and teacher ratings of ODD dimensions and emotion regulation: informant discrepancies in a two-phase study. Journal of Child and Family Studies, 1-11. https://doi.org/10.1007/s10826-021-02168-y

Mitsis, E. M., McKAY, K. E., Schulz, K. P., Newcorn, J. H., & Halperin, J. M. (2000). Parent–teacher concordance for DSM-IV attention-deficit/hyperactivity disorder in a clinic-referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(3), 308-313. https://doi.org/10.1097/00004583-200003000-00012

Mulraney, M., Arrondo, G., Musullulu, H., Iturmendi-Sabater, I., Cortese, S., Westwood, S. J., ... & Coghill, D. (2022). Systematic review and meta-analysis: screening tools for attention-deficit/hyperactivity disorder in children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *61*(8), 982-996. https://doi.org/10.1016/j.jaac.2021.11.031

Murray, A. L., Booth, T., Ribeaud, D., & Eisner, M. (2018). Disagreeing about development: An analysis of parent-teacher agreement in ADHD symptom trajectories across the elementary school years. *International journal of methods in psychiatric research*, *27*(3), e1723. https://doi.org/10.1002/mpr.1723

Murray, D. W., Kollins, S. H., Hardy, K. K., Abikoff, H. B., Swanson, J. M., Cunningham, C., ... & Chuang, S. Z. (2007). Parent versus teacher ratings of attention-deficit/hyperactivity disorder symptoms in the Preschoolers with Attention-Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of child and adolescent psychopharmacology*, *17*(5), 605-619. https://doi.org/10.1089/cap.2007.0060

Nanda, A., Janga, L. S. N., Sambe, H. G., Yasir, M., Man, R. K., Gogikar, A., & Mohammed, L. (2023). Adverse effects of stimulant interventions for attention deficit hyperactivity disorder (ADHD): a comprehensive systematic review. *Cureus*, *15*(9). https://doi.org/10.7759/cureus.45995

Narad, M. E., Garner, A. A., Peugh, J. L., Tamm, L., Antonini, T. N., Kingery, K. M., ... & Epstein, J. N. (2015). Parent–teacher agreement on ADHD symptoms across development. *Psychological assessment*, *27*(1), 239. https://doi.org/10.1037/a0037864

National Research Council (US) Panel to Review the Status of Basic Research on School-Age Children, & Collins, W. A. (Eds.). (1984). *Development during middle childhood: The years from six to twelve*. National Academies Press. https://doi.org/10.17226/56

Olson, S. L., Davis-Kean, P., Chen, M., Lansford, J. E., Bates, J. E., Pettit, G. S., & Dodge, K. A. (2018). Mapping the growth of heterogeneous forms of externalizing problem behavior between early childhood and adolescence: A comparison of parent and teacher ratings. *Journal of abnormal child psychology*, *46*(5), 935-950. https://doi.org/10.1007/s10802-018-0407-9

Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical child and family psychology review*, *10*(4), 335-351. https://doi.org/10.1007/s10567-007-0027-3

Owens, J. A., Maxim, R., Nobile, C., McGuinn, M., & Msall, M. (2000a). Parental and self-report of sleep in children with attention-deficit/hyperactivity disorder. *Archives of*

pediatrics & adolescent medicine, *154*(6), 549-555. https://doi.org/10.1001/archpedi.154.6.549

Owens, J., Sangal, R. B., Sutton, V. K., Bakken, R., Allen, A. J., & Kelsey, D. (2009). Subjective and objective measures of sleep in children with attention-deficit/hyperactivity disorder. *Sleep medicine*, *10*(4), 446-456. https://doi.org/10.1016/j.sleep.2008.03.013

Owens, J. A., Spirito, A., McGuinn, M., & Nobile, C. (2000b). Sleep habits and sleep disturbance in elementary school-aged children. *Journal of Developmental and Behavioral Pediatrics*, *21*(1), 27-36. https://doi.org/10.1097/00004703-200002000-00005

Peterson, B. S., Trampush, J., Brown, M., Maglione, M., Bolshakova, M., Rozelle, M., ... & Hempel, S. (2024a). Tools for the diagnosis of ADHD in children and adolescents: a systematic review. *Pediatrics*, *153*(4). https://doi.org/10.1542/peds.2024-065854

Peterson, B. S., Trampush, J., Maglione, M., Bolshakova, M., Rozelle, M., Miles, J., ... & Hempel, S. (2024b). Treatments for ADHD in Children and Adolescents: A Systematic Review. *Pediatrics*, *153*(4), e2024065787. https://doi.org/10.1542/peds.2024-065787

Petric, E., & Szamosközi, I. (2018). Parent-teacher agreement on oppositional defiant disorder symptoms: A meta-analysis. *Transylvanian Journal of Psychology*, *19*(1), 47–73. epsz.psiedu.ubbcluj.ro

Quinn, P. O., & Madhoo, M. (2014). A review of attention-deficit/hyperactivity disorder in women and girls: uncovering this hidden diagnosis. *Prim Care Companion CNS Disord*, *16*(3), PCC-13r01596. https://doi.org/10.4088/pcc.13r01596

Riddell, C., Nikolić, M., Dusseldorp, E., & Kret, M. E. (2024). Age-related changes in emotion recognition across childhood: A meta-analytic review. *Psychological Bulletin*, *150*(9), 1094. https://doi.org/10.1037/bul0000442

Rucklidge, J. J. (2010). Gender differences in attention-deficit/hyperactivity disorder. *Psychiatric Clinics*, *33*(2), 357-373. https://doi.org/10.1016/j.psc.2010.01.006

Rucklidge, J. J., Eggleston, M. J., Johnstone, J. M., Darling, K., & Frampton, C. M. (2018). Vitamin-mineral treatment improves aggression and emotional regulation in children with ADHD: a fully blinded, randomized, placebo-controlled trial. *Journal of Child Psychology and Psychiatry*, *59*(3), 232-246. https://doi.org/10.1111/jcpp.12817

Rucklidge, J. J., Frampton, C. M., Gorman, B., & Boggis, A. (2014). Vitamin–mineral treatment of attention-deficit hyperactivity disorder in adults: double-blind randomised placebo-controlled trial. *The British Journal of Psychiatry*, *204*(4), 306-315. https://doi.org/10.1192/bjp.bp.113.132126

Rucklidge, J., Taylor, M., & Whitehead, K. (2011). Effect of micronutrients on behavior and mood in adults with ADHD: Evidence from an 8-week open label trial with natural extension. *Journal of attention disorders*, *15*(1), 79-91. https://doi.org/10.1177/1087054709356173

Salari, N., Ghasemi, H., Abdoli, N., Rahmani, A., Shiri, M. H., Hashemian, A. H., ... & Mohammadi, M. (2023). The global prevalence of ADHD in children and adolescents: a systematic review and meta-analysis. *Italian journal of pediatrics*, *49*(1), 48. https://doi.org/10.1186/s13052-023-01456-1

Sayal, K., & Goodman, R. (2009). Do parental reports of child hyperkinetic disorder symptoms at school predict teacher ratings? *European child & adolescent psychiatry*, *18*(6), 336-344. https://doi.org/10.1007/s00787-009-0735-y

Sciberras, E., Efron, D., & Iser, A. (2011). The child's experience of ADHD. *Journal of Attention Disorders*, *15*(4), 321-327. https://doi.org/10.1177/1087054710361671

Sciberras, E., Roos, L. E., & Efron, D. (2009). Review of prospective longitudinal studies of children with ADHD: Mental health, educational, and social outcomes. *Current Attention Disorders Reports*, *1*(4), 171-177. https://doi.org/10.1007/s12618-009-0024-1

Sollie, H., Larsson, B., & Mørch, W. T. (2013). Comparison of mother, father, and teacher reports of ADHD core symptoms in a sample of child psychiatric outpatients. *Journal of Attention Disorders*, *17*(8), 699-710. https://doi.org/10.1177/1087054711436010

Sonuga-Barke, E. J., Koerting, J., Smith, E., McCann, D. C., & Thompson, M. (2011). Early detection and intervention for attention-deficit/hyperactivity disorder. *Expert review of neurotherapeutics*, *11*(4), 557-563. https://doi.org/10.1586/ern.11.39

Sprafkin, J., Gadow, K. D., Salisbury, H., Schneider, J., & Loney, J. (2002). Further evidence of reliability and validity of the Child Symptom Inventory-4: parent checklist in clinically referred boys. *Journal of Clinical Child and Adolescent Psychology*, *31*(4), 513-524. https://doi.org/10.1207/s15374424jccp3104_10

Staff, A. I., Oosterlaan, J., van der Oord, S., Hoekstra, P. J., Vertessen, K., de Vries, R., ... & Luman, M. (2021). The validity of teacher rating scales for the assessment of ADHD symptoms in the classroom: A systematic review and meta-analysis. *Journal of attention disorders*, *25*(11), 1578-1593. https://doi.org/10.1177/1087054720916839

Szklo, M., & Nieto, F. J. (2014). *Epidemiology: beyond the basics*. Jones & Bartlett Publishers.

Tamm, L., Loren, R. E., Peugh, J., & Ciesielski, H. A. (2021). The association of executive functioning with academic, behavior, and social performance ratings in children with ADHD. *Journal of Learning Disabilities*, *54*(2), 124-138. https://doi.org/10.1177/0022219420961338

Underwood, M. K. (1997). Peer social status and children's understanding of the expression and control of positive and negative emotions. *Merrill-Palmer Quarterly (1982-)*, 610-634. http://www.jstor.org/stable/23093362

Varni, J. W., Magnus, B., Stucky, B. D., Liu, Y., Quinn, H., Thissen, D., ... & DeWalt, D. A. (2014). Psychometric properties of the PROMIS® pediatric scales: precision, stability, and comparison of different scoring and administration options. *Quality of Life Research*, *23*(4), 1233-1243. https://doi.org/10.1007/s11136-013-0544-0

Varni, J. W., Thissen, D., Stucky, B. D., Liu, Y., Magnus, B., He, J., ... & DeWalt, D. A. (2015). Item-level informant discrepancies between children and their parents on the PROMIS® pediatric scales. *Quality of life Research*, *24*(8), 1921-1937. https://doi.org/10.1007/s11136-014-0914-2

Wolraich, M. L., Lambert, E. W., Bickman, L., Simmons, T., Doffing, M. A., & Worley, K. A. (2004). Assessing the impact of parent and teacher agreement on diagnosing attention-deficit hyperactivity disorder. Journal of Developmental & Behavioral Pediatrics, 25(1), 41-47. https://doi.org/10.1097/00004703-200402000-00007

World Health Organization. (2019). *International classification of diseases for mortality and morbidity statistics (11th ed.)*. https://icd.who.int/