BROADENING CONSTRUCT VALIDITY OF THE STRENGTHS AND DIFFICULTIES HYPERACTIVITY-INATTENTION SUBSCALE AMONG DIVERSE ETHNIC CHILDREN

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF THE ARTS

IN

PSYCHOLOGY

DECEMBER 2024

By:

Ashlyn W. W. A. Wong

Thesis Committee:

Yiyuan Xu, Chairperson

Brad Nakamura

Patrick Goh

Abstract

The current study aimed to further explore the construct validity of a broad-band, screening measure—Strengths and Difficulties Questionnaire Hyperactivity-Inattention Subscale (SDQ-HI)—by first analyzing its factor structure and subsequently assessing for measurement invariance and latent mean differences. The study sample consisted of 6,673 children, aged 4-17, from the 2019 National Health Interview Survey Sample Child Files, inclusive of White, Black, Hispanic/Latinx, and Asian children. A series of confirmatory factor analyses were used to establish the factor structure of the SDQ-HI and determine configural invariance across ethnic groups. Furthermore, the alignment optimization approach was employed to examine subsequent measurement invariance of the SDQ-HI and compare latent means between the groups. Results from this study indicated that 1) the SDQ-HI best fit a two-factor structure, 2) the SDQ-HI was deemed invariant across White, Black, Hispanic/Latinx, and Asian groups, except for scalar invariance, and 3) no significant differences were found in latent means between White and ethnically minoritized groups. The findings from this study help to support the use of the SDQ-HI as a screening tool across ethnic groups and provide important implications for considering the cultural impact toward behavior perceptions and endorsements. Limitations and future research directions are discussed.

Table of Contents

Broadening Construct Validity of the Strengths and Difficulties Hyperactivity-Inattention Subscale among Diverse Ethnic Children

Background and Purpose

In the United States (U.S.), Attention-Deficit/Hyperactivity Disorder (ADHD) is one of the most highly prevalent mental health disorders diagnosed among children (Bitsko et al., 2022; Centers for Disease Control and Prevention, CDC, 2022). Approximately 9.8% of children (6 million), aged 3-17, have received an ADHD diagnosis, with studies projecting an upward trend in diagnoses (CDC, 2022; Fairman et al., 2020; Getahun et al., 2013; Kiely & Adesman, 2015). Further, there appears to be an increase in the severity of ADHD symptoms following the COVID-19 pandemic (Bozinovic et al., 2021; Rogers & MacLean, 2023).

According to the *Diagnostic and Statistical Manual of Mental Health Disorders* (5th edition, text revision; DSM-V-TR), ADHD is categorized as a neurodevelopmental disorder with symptoms including persistent patterns of hyperactivity-impulsivity and/or inattention that significantly impacts one's functioning across two or more settings (American Psychiatric Association; APA, 2022). The three most recent changes in the DSM-V-TR criteria for ADHD include (1) changing the age of onset to age 12 instead of 7, (2) defining ADHD in terms of symptoms rather than impairment, and (3) substituting *subtypes* with *predominant presentations* (i.e., Epstein & Loren, 2013; Koutsoklenis & Honkasilta, 2022). Although there is a lack of understanding regarding causal mechanisms that may contribute toward the increase in ADHD prevalence, some factors may involve greater recognition and awareness of ADHD and mental health, along with the changes made to the DSM over the years that have made it easier to meet criteria (Getahun et al., 2013).

The reported increase in ADHD prevalence rates is cumulative among children within the U.S.; however, when examining prevalence rates by different sociodemographic subgroups, disparities in ADHD prevalence become apparent. Disparities within ADHD prevalence have been widely documented regarding sex differences: boys have higher ADHD prevalence rates than girls (APA, 2022; Fairman et al., 2020; Getahun et al., 2013). Further, boys are more likely to display externalizing, hyperactive symptoms than girls who are more likely to present with inattention symptoms (Gaub & Carlson, 1997; Loyer Carbonneau et al., 2021). Some theories, or explanations, proposed by scholars to understand these patterns include gendered socialization/stereotypes, hormonal activation differences, higher thresholds in girls, greater symptom severity in boys, and referral biases (Arnett et al., 2015; Lynch & Davidson, 2022; Martel, 2013; Quinn & Madhoo, 2014; Rhee et al., 1999; Rucklidge, 2010). However, another disparity that has received less attention surrounds culture, race, and ethnicity. The extant literature regarding this issue has consistently found that children of racially and ethnically minoritized groups (i.e., Black, Hispanic/Latinx, Asian, Native Hawaiian/Pacific Islander, American Indian, Alaska Native) are less likely to be diagnosed and have lower ADHD prevalence rates in comparison to their White counterparts (Chung et al., 2019; Coker et al., 2016; Davis et al., 2021; Fairman et al., 2020; Getahun et al., 2013; Shi et al., 2021; Wong & Landes, 2022). Additionally, racially and ethnically minoritized groups are less likely to receive treatment (e.g., medication, behavioral therapy, psychosocial therapy) despite being diagnosed with ADHD (Bauermerister et al., 2003; Davis et al., 2021; Yang et al., 2022). However, unlike research regarding sex differences in ADHD, there is a lack of understanding regarding racial and ethnic differences in ADHD prevalence.

Further research regarding racial and ethnic differences in ADHD identification is warranted given that both short- and long-term outcomes may differ depending on whether one receives appropriate treatment and support. When children are left undiagnosed and subsequently untreated, despite being presented with clinically elevated ADHD-related symptoms, this can lead to a variety of negative outcomes that persist into adulthood (Quinn & Madhoo, 2014). In general, untreated ADHD symptoms may continue to worsen over time, and this could evolve into difficulties with academic, social, and health functioning (Bussing et al., 2010; Quinn & Madhoo, 2014). For example, ADHD is associated with lower levels of academic achievement, greater peer and socialization problems, and an increased likelihood of participating in risky behaviors, such as substance use and sexual behaviors (APA, 2022; Fleming, 2017). Further, when children get older, this can lead to higher rates of unemployment, job instability, incarceration, physical health difficulties (e.g., obesity, hypertension), premature death, and the development of other comorbid mental health disorders (e.g., schizophrenia, conduct disorder, substance use disorders; APA, 2022; Barbaresi et al., 2013; Bussing et al., 2010; London & Landes, 2016; Loyer Carbonneau et al., 2021; Morgan et al., 2016). Identification and treatment of ADHD at a young age (i.e., between 6 to 8 years of age) plays a pivotal role in improving outcomes and mitigating negative consequences associated with ADHD (Zahmacioglu & Kilic, 2017). Thus, having measures that can adequately screen and assess ADHD across racial and ethnic groups is imperative.

Despite the importance of early identification, there are several barriers and mediating mechanisms that may inhibit racially and ethnically minoritized groups from receiving appropriate identification and care, both systemic and cultural. Language differences, bias, measurement variance, and reduced access to mental health services, as well as stigmatization toward mental health and higher distress thresholds in racially and ethnically minoritized families, may impede the identification of ADHD (Cook et al., 2017; Liang et al., 2016; Hays et al., 2010). It remains largely unclear what the primary inhibitor is, or if it is an accumulation of these factors, as there is little-known work in this area of research. However, the first step to providing greater understanding is to ensure that ADHD measures are psychometrically sound and culturally sensitive.

The purpose of my proposed study was to further analyze the construct validity of a widely used, broad-band rating scale in hopes of determining whether the measure can be used as an appropriate screening tool for identifying racially and ethnically minoritized children struggling with ADHD symptoms. More specifically, I examined the factor structure and assessed for measurement invariance of the Strengths and Difficulties Questionnaire Hyperactivity-Inattention Subscale (SDQ-HI) among White, Black, Hispanic/Latinx, and Asian children in a representative sample of the U.S. population from the 2019 National Health Interview Survey (NHIS). After determining whether the SDQ-HI was invariant across these groups, I analyzed whether ratings varied as a function of the child's racial and ethnic background by examining latent mean differences.

A Working Definition of Race and Ethnicity

The terms 'race' and 'ethnicity' are used interchangeably in research, and there are multiple ways in which they are defined. In alignment with the Office of Management and Budget (OMB, 1997), which establishes racial and ethnic categories for the U.S. Census, the NHIS collects race and Hispanic/Latinx ethnicity data from two separate questions and combines these questions to create a race and ethnicity variable[1]. Race and ethnicity are not distinctly

[1] OMB (2024) has recently revised its standards to include Hispanic/Latinx into a single, combined race and ethnicity question. Given that the current study uses data from the 2019 NHIS, we will be using previous standards.

defined by the NHIS and most definitions from the OMB indicate that both race and Hispanic/Latinx ethnicity indicate a person's origins in different geographic regions (NCHS, 2020; OMB, 1997). Given the ambiguity of these definitions and that a combined racial and ethnic term is not recommended due to its confusing implications (Flanagin et al., 2021; Lu et al., 2022), I will be using the term 'ethnicity' in isolation moving forward. Race is typically a socially defined construct, where categories are made based on physical traits and geographic origin (Lu et al., 2022; Phinney, 1996). In this respect, ethnicity will be used to encapsulate race, as it can be defined as a construct that incorporates both cultural characteristics of a group (e.g., language, values, practices), as well as their ancestry and origin from geographic regions (Lu et al., 2022; Phinney, 1996).

Ethnic Disparities in ADHD Identification

Within the U.S., disparities in equitable health care are ubiquitous across ethnic groups. Compared to the dominant White group, ethnically minoritized groups (e.g., Hispanic/Latinx, Black, Asian) disproportionately receive physical and mental health care, and when they do, they are more likely to obtain lower-quality services (Fiscella et al., 2000; Lau et al., 2012; Weinstein et al., 2017). Consequently, ethnically minoritized groups have worse health outcomes, as well as greater morbidity and mortality rates than their White counterparts (Weinstein et al., 2017; Xu et al., 2022). In the context of mental health, ethnically minoritized adults are less likely than White adults to be diagnosed with a mental health disorder yet may have an elevated risk for chronic persistence of disorders throughout their developmental life span (Jimenez et al., 2010; Vilsaint et al., 2019). Further, in a nationally representative sample of over 20,000 adults, aged 18 and older, White adults had the highest lifetime prevalence rate of a mental disorder (45.6%), while Asian adults had the lowest rate (23.5%) followed by Black (37.0%) and Hispanic/Latinx

(38.8%) adults (Alvarez et al., 2019). These patterns are also evident among children, as ethnically minoritized children are less likely to be diagnosed with and have a lower prevalence of mental health disorders (e.g., depression, anxiety, substance use disorders, autism spectrum disorder) than White children (Bitsko et al., 2018; Mak & Rosenblatt, 2002; Nguyen et al., 2007).

The lower prevalence of mental health disorders among Hispanic/Latinx, Black, and Asian groups compared to their White counterparts could be potentially related to several factors. This includes ethnically minoritized adults having less access to and utilization of mental health services, as well as less education about mental health contributing to stigmatized beliefs, compared to White adults (Ault-Brutus & Alegria, 2016; Cook et al., 2017). Additionally, these patterns could be partially related to culturally differential experiences, coping abilities, and support systems. Prior studies have found that White adults with lower levels of education and income may have an increased risk of exposure to stressors, such as discrimination from disparate expectations and realities, and lack strategies/resources to effectively cope with their circumstances (Vilsaint et al., 2019; Bower et al., 2013). Comparatively, other studies have suggested that ethnically minoritized adults, who experience discrimination more frequently and have been impacted across generations, might be more prepared to deal with exposure to stressors (i.e., culturally resilient, Spence et al., 2016).

Health inequities in ethnically minoritized adults are well documented throughout the literature. However, an emergent area requiring further exploration surrounds efforts in identifying mental health disparities among ethnically minoritized children, particularly regarding ADHD. The current literature documenting prevalence rates and likelihood ratios of ADHD has sustained that Hispanic/Latinx, Black, and Asian children have lower rates and are less likely to be diagnosed than White children. However, like the discourse on ethnic differences in mental health, it is unclear whether results reflect underestimations, overestimations, or *true* differences across ethnicity.

ADHD Prevalence Across Ethnic Groups

Within multiple epidemiological studies and reports of diagnostic prevalence in the U.S. across time, there is a consistent finding that ethnically minoritized children have lower odds and rates of being diagnosed with ADHD in comparison to White children (Boyle et al., 2011; Coker et al., 2016; Fairman et al., 2017; Morgan et al., 2013; Pastor & Reuben, 2005). In a nationally representative, longitudinal study of 17,100 kindergarten children from 1998, Black children were 69% and Hispanic/Latinx children were 50% less likely than White children to be diagnosed with ADHD even while controlling for potential confounds (e.g., gender, age, socioeconomic status [SES], health insurance status, geographic region, parent age, academic achievement, language; Morgan et al., 2013). Further, among a sample of 4,297 parent-child pairs of children surveyed longitudinally across fifth- (2004-2006), and tenth-grade (2009-2011), Black (.40 for fifth-grade and .42 for tenth-grade) and Hispanic/Latinx children (.37 and .46) had significantly lower adjusted odds of ever receiving an ADHD diagnosis in comparison to White children (Coker et al., 2016). Prevalence rates have also been shown to be lower among ethnically minoritized groups using data from the National Health Interview Survey (NHIS), which collects nationally representative health and sociodemographic data from children and adults. Pastor and Reuben (2005) analyzed data from the 1997-2001 NHIS and found that in a sample of 21,294 children, aged 6-11, Black (2.2%) and Hispanic/Latinx children (1.4%) had a lower prevalence of ADHD than White children (4.4%). Similarly, Boyle and colleagues (2011) analyzed aggregate data from the 1992-2008 NHIS with a sample of 119,367 children, aged 3-

17, and reported that Hispanic/Latinx children had significantly lower rates of ADHD (3.9%) than White (7.8%) and Black children (6.3%).

The literature on this subject focused primarily on differences among Black, Hispanic/Latinx, and White children; however, recently there has been more research including findings on Asian children. Additionally, recent literature suggests that Black children have a higher prevalence of ADHD than previously reported (Cénat et al., 2022). In a study using 842,830 medical records from a California state system of care across ten years (i.e., 2001-2010), Asian and Pacific Islander (1.2% in 2001 and 1.2% in 2010) children and children of other ethnic groups (3.1% and 3.4%) had lower prevalence of ADHD than White (4.7% and 5.6%), Black (2.6% and 4.1%), and Hispanic/Latinx children (1.7% and 2.5%; Getahun et al., 2013). Additionally, cross-sectional studies provided similar evidence. Reyes and colleagues (2013) examined data from 7,954 children, aged 5-17, in Dane County, Wisconsin medical clinics from 2007-2008. They found that Black children (6.8%) had the highest prevalence of ADHD followed by White (4%), Native American (2.7%), Hispanic (1.6%), and Asian (1.3%) children. Further, Davis and colleagues (2021) found that in their cross-sectional sample of 693,942 children (ages 6-17) receiving Kentucky Medicaid in 2017, Asian (1%) and Hispanic/Latinx (6%) children had significantly lower ADHD prevalence rates compared to Black (14%) and White (15%) children.

Although these studies are limited in their generalizability, other studies that used nationwide datasets corroborate the above findings. A study by Shi and colleagues (2021) examined data from a national insurance database, with 283,011 children (approximate mean age of 6 years) born between 2006 and 2012 and found that Asian, Hispanic/Latinx, and Black children were 52%, 23%, and 17% less likely to be diagnosed with ADHD compared to White children, respectively. Furthermore, Wong and Landes (2022) used a sample of 120,129 children, aged 5-17 from the 2004-2018 NHIS. The authors found that Asian (81%) and Hispanic/Latinx (47%) children were significantly less likely than White children to obtain a diagnosis of ADHD while accounting for age and the survey year. On the other hand, Black children had similar odds with White children of receiving an ADHD diagnosis (Wong & Landes, 2022).

Summary and Critique

Given findings from these large epidemiological studies and reports, there is robust evidence that the prevalence of ADHD differs based upon the child's ethnic background, with Hispanic/Latinx and Asian children having lower rates and White and Black children having higher rates. Despite an expansion of the literature, with the inclusion of Asian child populations, there is still a lack of understanding as to *why* ADHD prevalence rates and likelihood ratios vary significantly across ethnic groups. A limitation of studies examining prevalence rates and likelihood ratios is that they end up being descriptive rather than explanatory. Regardless, there are several ways that researchers attempted to interpret their findings.

Some researchers suggested that White children had higher prevalence and symptom ratings of ADHD than ethnically minoritized children due to having higher socioeconomic status (SES), larger odds of parental referrals to receive support for schooling (e.g., stimulant medications, special education services), and greater demand from parents to diagnosis behaviors as ADHD-related (APA, 2022; Getahun et al., 2013; Owens, 2021). Other researchers discussed that ethnically minoritized groups may have lower prevalence and symptom ratings due to a mixture of cultural issues and systemic barriers (Cenat et al., 2022; Pastor & Reuben, 2005; Wong & Landes, 2022). Thus, some ethnic groups might be overestimated while others are underestimated. Furthermore, an alternative interpretation could be that White children may have a higher susceptibility for ADHD than ethnically minoritized groups, and these differences reflect true ethnic variation in the presentation of ADHD. However, due to the complexity behind these disparities and studies not having the data to explore potential contributing factors, there remains a lack of theoretical interpretation.

Nonetheless, these studies are still imperative, given that they raise awareness about the issue of ethnic disparities in childhood ADHD, and new findings provide increased attention to potential areas needing further exploration. For example, the recent finding that Black children have a higher prevalence and likelihood of obtaining an ADHD diagnosis than White children and children of other ethnic groups contradicts earlier findings (Cénat et al., 2022; Davis et al., 2021; Getahun et al., 2013; Reyes et al., 2013; Wong & Landes, 2022). This may be indicative that Black children may have been underdiagnosed despite having symptoms of ADHD (Coker et al., 2016), and this may be a similar case for other ethnic minority children. Given this new finding, as well as the consistency of ethnic disparities found across ADHD prevalence and likelihood ratios, the demand for efficient and culturally sensitive measures identifying ADHD has risen (Hays et al., 2010). The current study will help to address ethnic differences in ADHD identification by examining the construct validity of a clinical screening tool in reporting symptoms of ADHD across ethnic groups.

Issues and Considerations in Using ADHD Rating Scales across Ethnic Groups

The process for identifying ADHD in children typically begins with a referral from either the school or a medical professional (Bussing et al., 1998). Additionally, parents may refer their child for evaluation if they have not been referred by other sources. To receive an ADHD diagnosis, children must undergo an assessment with a mental health professional (e.g., psychologist, psychiatrist, primary care provider). The assessment consists of diagnostic interviews with the child and their parent/legal guardian, behavioral observations, and a physical examination, if conducted in a primary care setting (Langberg et al., 2008). Additionally, according to the American Academy of Pediatrics (AAP), there is a set of guidelines to assist clinicians in the assessment and diagnosis of ADHD: (1) The clinician should evaluate any child who presents with difficulties in hyperactivity, impulsivity, and inattention, (2) determine that the child meets DSM criteria for ADHD and rule out other disorders, and (3) screen for comorbid conditions (Wolraich et al., 2019). To satisfy guideline two, there is a need to document symptoms and impairment in two or more settings through a combination of parent and teacher reports, as well as self-report (APA, 2022; Langberg et al., 2008; Wolraich et al., 2019). Therefore, part of the assessment procedure can be facilitated using ADHD rating scales.

Rating scales are not a replacement for diagnostic interviews and cannot be used independently to provide a diagnosis (Bussing et al., 1998). Instead, these measures are used as a *screening tool* in the assessment process to help focus attention on different problem areas, increasing the likelihood that an individual refers a child for assessment (Langberg et al., 2008). Rating scales can be used by multiple informants, as many offer different versions (i.e., parent report, teacher report, child report). Parents, teachers, and sometimes the child themselves, are asked to fill out these rating scales, with higher scores indicating a greater frequency of symptoms and behaviors. The various reporting forms offer further insight into the child's behaviors in multiple environments. Additionally, the utilization of adult rather than child, self-reports increases the accuracy of diagnosis, since adult informants are more reliable than children when assessing for externalizing behaviors like those described in ADHD (Collett et al., 2003).

For this study, I will be focusing on adult reports of rating scales and shall not be discussing child, self-report versions of ADHD rating scales.

Types of ADHD Rating Scales

There are many types of rating scales available, both broad- and narrow-based. Some well-established measures include the Child and Behavior Checklist (CBCL; Achenbach & Rescorla, 2000), Behavioral Assessment System for Children (BASC; Reynolds & Kamphaus, 1992), Children's Symptom Inventory – 4th Edition (CSI-4; Gadow & Sprafkin, 1994), and Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). The CBCL, BASC, and SDQ are known as *broad-band* rating scales, which help to screen for multiple psychopathologies (e.g., anxiety, depression, oppositionality, hyperactivity), offering a holistic approach to evaluating children (Collett et al., 2003; Langberg et al., 2008; Pelham et al., 2005). These broad-band scales are useful, given that ADHD is often comorbid and may be differentially diagnosed with other disorders (APA, 2022). On the other hand, broad-band scales may not offer a full set of symptomatology for ADHD, as there are fewer items per psychopathology domain, and items are typically empirically and rationally derived (Pelham et al., 2005). Therefore, these scales are limited in their comprehensiveness. Other rating scales include the Conners' Rating Scale (CSR; Conners, 1997), the Swanson, Nolan and Pelham-IV Questionnaire (SNAP-IV; Swanson, 1992), Vanderbilt ADHD Diagnostic Rating Scales (VADRS; Wolraich et al., 1998), ADHD Rating Scale (ADHD-RS; DuPaul, Power, et al., 1998), and Hyperactivity Rating Scale (HRS; Spring et al., 1977). The CSR, SNAP-IV, VADRS, ADHD-RS, and HRS are known as *narrow-band* rating scales, which specifically assess for ADHD and are typically based on symptomatology in the DSM (Collett et al., 2003). Thus, narrow-band rating scales offer a more robust examination of ADHD while forfeiting comparison with other psychopathology.

Clinical Utility and Effectiveness of ADHD Rating Scales

Although broad- and narrow-band rating scales have different pros and cons, they are both clinically useful and effective. Studies analyzing psychometric properties of broad- and narrow-band scales reveal that they are well-validated and have solid reliability (Collett et al., 2003; Pelham et al., 2005). In addition, many of the broad-band rating scales are highly correlated with one another, despite consisting of empirically and rationally derived items (Pelham et al. 2003). Furthermore, meta-analyses have shown that ADHD rating scales maintain moderate specificity and sensitivity rates in detecting ADHD (i.e., CBCL, CRS; Chang et al., 2016), as well as excellent overall diagnostic accuracy (Mulraney et al., 2022). Therefore, both narrow- and broad-based scales are acceptable options for use, and implementing them in tandem is optimal for enhancing the assessment process (Collett et al., 2003).

Ethnic Differences in ADHD Ratings

Despite evidence supporting that ADHD rating scales are effective and clinically useful tools in the ADHD assessment process, research looking at ADHD ratings across ethnic groups indicates potential limitations to using ADHD measures universally as a screening tool. More specifically, average ratings tend to vary by the child’s ethnic background, raising the question of whether these mean differences were due to nonequivalent measurement properties across ethnic groups, which have been rarely examined in the literature.

Ratings for Black Children. Older studies using community samples of children from schools indicate that Black children are rated higher on ADHD rating scales (i.e., ADHD-RS-IV, CRS, HRS, SNAP-IV) than White children by their teachers (Bussing et al., 2008; DuPaul et al., 1997; Langsdorf et al., 1979; Spring et al., 1977). Villodas and colleagues (2019) randomly selected 314 students (mean age of 8 years) from 17 elementary schools in a large, public school district in the 1994-1995 school year. They found that Black children were rated by teachers significantly higher on the inattention symptoms scale in the CSI-4 than White children. Similarly, in a small meta-analysis of five studies (Arnold et al., 2003; Reid et al., 2001; DuPaul, Anastopoulos, 1998; Epstein et al., 2005; Reid et al., 1998), Miller and colleagues (2009) found a significant pooled effect, with medium effect size, indicating that Black children had higher teacher ratings of ADHD (i.e., ADHD-RS-IV, CRS, SNAP-IV) than White children. Although one study indicated that parents rated Black children higher than White children (DuPaul, Anastopoulos, et al., 1998), four out of the five studies relied on teacher-report rather than parent-report of ADHD rating scales (Miller et al., 2009). A recent study by Evans and colleagues (2013) looking at 875 children, aged 4-18, and 143 teachers from four states (e.g., Kansas, Maryland, Ohio, Pennsylvania) further substantiates prior findings. The authors found that teacher scores on the inattention and hyperactivity-impulsivity scales on the Disruptive Behavior Disorders Rating Scale (DBD) were significantly higher for Black children than White children.

Other studies indicate that parent ADHD ratings among Black children are lower than among White children. Hillemeier and colleagues (2007) found that in a longitudinal sample of 1,070 children from the Fast Track Project—a project looking at risk for emotional/behavioral among low SES neighborhoods in four states (i.e., North Carolina, Tennessee, Pennsylvania, Washington) —African American children had higher parent endorsements of ADHD symptoms on the Diagnostic Interview Schedule for Children (DISC) than White children. However, parents of African American children had a higher likelihood of endorsing ADHD symptoms than parents of White children when they were balanced on the same hyperactivity levels (Hillemeier et al., 2007). Although DISC is a structured interview tool, other ADHD rating scales, such as the CRS and the VADRS, use similar ADHD symptom descriptions (Hillemeier et al., 2007); thus, using the DISC as a proxy for ADHD rating scales, this study indicates that parents of White and Black children may differentially report ADHD symptoms, despite similar display of behaviors. In a recent study, DuPaul and colleagues (2016) collected 2,079 parent and 1,509 teacher ratings of children, aged 5-17, from across the U.S. in 2014 using the ADHD-RS-V Home and School versions. Teachers rated Black children higher than White children for both inattention and hyperactive-impulsive scales while parents did not rate Black children higher on hyperactive-impulsive and inattention scales than White children (DuPaul et al., 2016). Furthermore, another study by Wexler and colleagues (2022) looked at a clinically referred sample of 3,943 children, aged 6-18, whose data was collected between 2011-2020 from a neuropsychology clinic. They further evidenced that while teachers endorsed greater hyperactive-impulsive symptoms on the ADHD-RS-V for Black children than White children, parents of Black children endorsed fewer symptoms than parents of White children.

Ratings for Hispanic/Latinx Children. As for Hispanic/Latinx children, the literature is mixed; however, results relatively suggest that Hispanic/Latinx children have lower ADHD ratings than White and Black children. Ramirez and Shapiro (1998) looked at CRS ratings from 61 White teachers regarding 61 Hispanic/Latinx and 61 White children, aged 6-11, from an urban, school district. Hispanic/Latinx children were rated by teachers similarly or lower on ADHD symptoms, but not higher than White children (Ramirez & Shapiro, 1998). Additionally, using 10,367 children, aged 4-17, from the 2001 NHIS, Cuffe and colleagues (2005) found that Hispanic/Latinx children had significantly lower ADHD symptom ratings than Black children on the parent-reported SDQ. In the study by DuPaul and colleagues (2016), teachers rated Hispanic/Latinx children lower on the hyperactive-impulsive and inattention subscales than

Black children and parents rated Hispanic/Latinx children lower than White children on inattention scales of the ADHD-RS-V.

Ratings for Asian Children. In terms of Asian children, studies consistently found lower ADHD ratings among parents and teachers when compared to White and Black children. Earlier studies show that within different ADHD rating scales (e.g., CBCL, HRS), parents and teachers of Asian children in community samples, endorsed lower ratings of ADHD symptoms than for White and Black children (Spring et al., 1977; Wells et al., 1992). In the study by Villodas and colleagues (2019) from the 1994-1995 school year, White and Black children were rated by teachers significantly higher than Asian children on the inattention symptoms scale in the CSI-4 (Villodas et al., 2019). Additionally, when looking at parent and teacher ratings together, Morgan and colleagues (2016) examined a sample of 7,456 children from kindergarten to 8th grade who participated in the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999. The authors obtained ratings of ADHD and comorbid conduct disorder (CD) symptomatology from parents and teachers, using scales they developed from DSM criteria, and found that Asian children had significantly lower adjusted odds of moderate (.4) and severe (.07) ADHD and CD symptomatology than White children.

Further, recent studies found supportive evidence of lower ratings among Asian children. In the study by DuPaul and colleagues (2016), teachers rated Asian children lower on hyperactive-impulsive and inattention subscales on the ADHD-RS-V than Black children. Meanwhile, Cuffe and colleagues (2005) found that children within the “other” ethnic category, which was comprised of Asian, American Indian, and Pacific Island children, had significantly lower ADHD symptom ratings than both Black and White children on the parent-reported SDQ (Cuffe et al., 2005). In addition, Quebles and colleagues (2020) analyzed data from 2,567 children with autism spectrum disorder (ASD), aged 6-18, who were registered in the Autism Treatment Network Registry between 2008 and 2016. Using the CBCL, they found that Asian children had significantly lower odds of hyperactivity (.49) than White children when rated by their parents.

Overall, the extant literature indicates that ethnicity may affect how informants rate ADHD-related behaviors and determine clinical significance. Studies including Black children have robust evidence that teachers are more likely to rate Black children higher than White children (DuPaul et al., 1997; Langsdorf et al., 1979; Miller et al., 2009; Spring et al., 1977). As for parents, the results are mixed. Although an older study indicated higher parent ratings (DuPaul, Anastopoulos, et al., 1998), recent studies suggest that Black parents are more likely than White parents to rate their child lower on ADHD scales (DuPaul et al., 2016; Hillemeier et al., 2007; Wexler et al., 2022). Studies involving Asian and Hispanic/Latinx children, on the other hand, indicate both lower parent and teacher ratings of ADHD than White and Black children (Cuffe et al., 2005; DuPaul et al., 2016). These results indicate the need to further investigate the utilization of ADHD rating scales in the context of multiple ethnic groups.

Limitations in Using ADHD Rating Scales across Ethnicity

There are several limitations when it comes to using ADHD rating scales, and these may inhibit the use of these measurements across ethnic groups, contributing to the differences found within ratings and subsequent diagnostic prevalence. These limitations include insufficient items assessing for functional impairment, language and linguistic inadequacies, and susceptibility to ethnic biases.

Lack of Measuring for Impairment. To meet the criteria for ADHD, clinicians must assess not only symptoms but also impairment in two or more settings (APA, 2022). Although symptoms and impairment can and should be further examined using diagnostic interviews, including measures of impairment within the rating scale can assist the diagnostic process. However, not all ADHD rating scales measure for impairment (Langberg et al., 2008; Pelham et al., 2005), which may de-emphasize the importance of considering functional impairment in assessment. Thus, other scales that are specific to impairment would need to be included, such as the Columbia Impairment Scale (Bird et al., 1993) or the Brief Impairment Scale (Bird et al., 2005). Doing so may add extra costs, such as increased time and energy toward completing screening measures. Further, if additional measures assessing for impairment are dismissed, this may lead to an inaccurate diagnosis of ADHD, with over or underestimations of the disorder.

Additionally, some research has indicated that the inclusion of impairment is imperative in assessing for ADHD, as it may be a more universal construct than symptomatology across cultural groups (Arcia & Fernandez, 1998; Bauermeister et al., 2005; Gerdes et al., 2013). A study by Haack and colleagues (2018) found that although Hispanic/Latinx parents reported fewer ADHD symptoms than non-Hispanic/Latinx parents, they tended to report higher levels of impairment. Hispanic/Latinx parents may perceive their child's problems as impairing but not relate them to symptomatology due to their conceptualizations of ADHD and mental health in general (Gerdes et al., 2013; Liang et al., 2016). For example, Hispanic/Latinx parents may not recognize symptoms of ADHD, in comparison to impairment, due to their values in *familismo* (i.e., interdependence with the family/community), *respecto* (i.e., respect and compliance with authority), and *spiritualismo* (i.e., spirituality and religiosity; Gerdes et al., 2013; Haack et al., 2018; Lawton et al., 2016). These values may help parents maintain an external locus of control—the belief that external forces (e.g., fate, peer influences, teacher discipline style) may be contributing toward their child's behaviors—which decreases the likelihood that

Hispanic/Latinx parents perceive hyperactive-impulsive and inattentive behaviors as problematic (Gerdes et al., 2013; Lawton et al., 2016). On the other hand, Hispanic/Latinx parents might report more impairment-related problems because it affects their cultural values more, such as respecting and getting along with other people (Haack et al., 2018).

Although the literature in this area mostly focuses on the Hispanic/Latinx context, Black and Asian parents are also less likely to adopt a biomedical view of ADHD (Liang et al., 2016). Bussing and colleagues (2003) found that Black parents reported fewer worries surrounding ADHD-related problems in school than White parents. Though this may be related to distrust toward the school system, it may also be related to how Black culture appreciates expressive movement and language (Davison & Ford, 2001; Neal et al., 2003). Thus, Black parents' cultural views may influence the belief that ADHD-related behaviors are a normative part of their child's development. As for Asian parents, collectivistic values may increase feelings of stigmatization when endorsing ADHD symptoms, as they perceive their child's behaviors as a reflection of their parenting and a deviation from their beliefs in social unity and harmony (Liang et al., 2016; Pham, 2013). Additionally, relaxed parenting styles earlier in childhood and the adoption of an internal locus of control—an idea that their child's behaviors are controllable with the ability to improve over time—may increase Asian parents' tolerance toward ADHD-related behaviors (Braje et al., 2012; Jose et al., 2000; Pham, 2013). Thus, Asian parents might only report problems when they reflect severe impairment, especially if it interferes with academic success and schooling which are an important part of Confucian and Hindu principles (Jose et al., 2000; Liang et al., 2016; Pham, 2013). Therefore, including impairment helps to capture problems that may be going on that might not be culturally reflected by ADHD symptomatology. Taking this into consideration, having a measure that includes both symptomatology and impairment can help to increase cultural sensitivity (DuPaul et al., 2016).

Language and Linguistic Inadequacies. Another weakness is that although some rating scales have been translated to multiple languages, such as the CBCL, SDQ, SNAP-IV, and ADHD-RS, other rating scales are only available in English and Spanish (e.g., BASC, CSI-4, CRS, VADRS). The 2019 U.S. Census Bureau report stated that about 20% of the U.S. population (67.8 million people) spoke a language other than English, and though Spanish was the most common non-English language, other frequently spoken non-English languages include Chinese, Tagalog, Vietnamese, and Arabic (Dietrich & Hernandez, 2022). Additionally, of the people who spoke Chinese, Tagalog, Vietnamese, and Arabic, the proportion of people who spoke English "less than very well" ranged from 30% to 52% (Dietrich & Hernandez, 2022). The low availability of rating scales in different languages limits certain ethnic groups, such as Asians, from using the scales to their fullest potential. Those who are not fluent in the available languages may not be able to understand what is written and might not answer accurately nor be able to utilize them completely. Further, when rating scales are available in different languages, there may be translation errors and some terms may not be translated equally to other languages even with the use of back-translation techniques (Reid et al., 1995). If scales are not linguistically accurate, their psychometric properties must be further evaluated before they can be deemed reliable and valid for use (Gerdes et al., 2013).

Bias and Differential Reporting. Other weaknesses include that ADHD rating scales may be susceptible to informant biases, as well as low cross-informant agreement, with differential ratings among informants (Des Los Reyes & Kazdin, 2005; Pelham et al., 2005).

This is further elucidated when ethnicity is involved (APA, 2022), as evidence shows that parents and teachers of Black and Hispanic/Latinx children differ in their reported scores.

A comparison between parent and teacher ratings of ADHD among Black children reveals significant discrepancies. Studies show that teachers tend to endorse greater ADHD ratings for Black children than their parents. In a sample of 258 3-year-old children, who were part of a longitudinal study on behavioral problems, Black parents were more likely than White parents to endorse lower hyperactivity and inattention ratings on the BASC than teachers (Harvey et al., 2013). Furthermore, Lawson and colleagues (2017) found similar findings in their two study samples of (1) 1,056 clinic-referred children, aged 4-7, being assessed for ADHD in an urban, northeastern hospital, and (2) 317 clinic- and school-referred children, aged 6-12, being evaluated for ADHD and whose data was retrieved from data collection centers. In both study samples, teachers reported higher levels of hyperactivity-impulsivity and inattention on the ADHD-RS-IV for Black compared to White children, but no differences were seen in parent reports (Lawson et al., 2017). Additionally, when ethnicity is accounted for among teacher informants, this further complicates findings. Kang and Harvey (2020) used a sample of 71 Black parents and 60 White teachers selected from across 42 states in the U.S. who rated the behaviors of White and Black children from video clips using the VADRS. They found that White teachers rated Black boys' behaviors higher on the VADRS than Black parents (Kang & Harvey, 2020).

As for Hispanic/Latinx children, Haack and colleagues (2018) found that among 159 children, 2nd to 5th grade, from 27 schools in the San Francisco Unified School District, parents of Hispanic/Latinx children had a higher likelihood than Non-Hispanic/Latinx parents to endorse lower CSI scores on ADHD symptomatology compared to teachers. However, Harvey and colleagues (2013) found that Hispanic/Latinx mothers were more likely than White mothers to endorse higher ratings of hyperactivity and attention ratings on the BASC than teachers. The authors noted that this discrepancy may be partially due to Hispanic/Latinx mothers having a significantly higher likelihood of depression, which is associated with endorsement of higher ADHD ratings; however, though associations were present, they did not reach statistical significance (Harvey et al., 2013). Despite the mixed results among Hispanic/Latinx children, this further corroborates that bias may be involved in completing ADHD rating scales and that results should be interpreted cautiously.

Differential ratings may be attributed to a multitude of factors, including behavior differences due to the environment (e.g., home vs. school) and differing perceptions or interpretations of items (Des Los Reyes & Kazdin, 2005; Liang et al., 2016). Additionally, informants may be influenced by personal expectations and biases about children from differing ethnic backgrounds (Des Los Reyes & Kazdin, 2005; Liang et al., 2016). Stereotypes about Black and Hispanic/Latinx children, especially males, include that they are anger-prone, hostile, and difficult, which stems from overrepresentation in the criminal justice system and robust achievement gaps (Bruton & Robles-Piña, 2009; Harvey et al., 2013; Mak & Rosenblatt, 2002; Neal et al., 2003; Villodas et al., 2019). These stereotypical characterizations may influence informants to believe that Black and Hispanic/Latinx children have more impairments and severe difficulties. Similarly, informants may be biased toward Asian children because of their low crime rates, association with nondeviant peers, and the "model minority" stereotype—the perception that Asian individuals are high-achieving, smart, hard-working, and well-disciplined (Braje et al., 2012; Jang, 2002; Lee, 1994). This schematic portrayal of Asian children being easy-going and disciplined may influence informants to overlook present difficulties, contributing to the lower ADHD ratings observed (Villodas et al., 2019; Spring et al., 1977). Differential reporting may not necessarily be detrimental since it can help one to understand the role of contextual factors influencing a child's behaviors, which has important implications in identifying appropriate interventions (Des Los Reyes et al., 2015). However, given that there may be potential biases affecting informant reports as well, it is imperative to include ratings from multiple informants to ensure a comprehensive assessment of ADHD and enhance the validity of the diagnosis (De Los Reyes et al., 2015; Mulraney et al., 2022).

Measurement Invariance of ADHD Rating Scales across Ethnicity

In addition to these weaknesses, there is another limitation to using ADHD rating scales with varying ethnic groups: ADHD rating scales may not maintain equivalent theoretical structures of ADHD across diverse populations, resulting in biased estimations. It is not clear whether the factor structure and constructs of hyperactivity-impulsivity and inattention within ADHD rating scales are the same across ethnic groups (Miller et al., 2009), and this is often frequently assumed without further testing. To assert equivalence across ethnic groups, measurement invariance must be attained (Gerdes et al., 2013; Miller et al., 2009).

Measurement invariance is defined as "equivalence of a construct across groups or measurement occasions" (Putnick & Bornstein, 2016, p. 1), and in the context of ethnicity, it can help determine whether measures are culturally valid. There are four steps to achieving measurement invariance, in which each subsequent step assumes the previous: (1) configural, (2) metric, (3) scalar, and (4) residual (Kline, 2016; Putnick & Bornstein, 2016). *Configural invariance*, which is the least restrictive level, indicates that the factor structure, with the same items in each factor, is replicable in each group. *Metric invariance* (also known as weak invariance) signifies that each factor is demonstrated in the same way, or on the same scale, across each group. *Scalar invariance* (also known as strong invariance) illustrates that groups have similar intercepts or thresholds, meaning that each group does not respond systematically higher or lower than the others. Lastly, *residual invariance* (also known as strict invariance) means that the error, or unique variance, for each item is equivalent across groups. Although residual invariance is the last level for reaching measurement invariance, researchers often do not require residual variance, as this level can be difficult to achieve, and it may be unreasonable to require this degree of precision in measurement (Kline, 2016; Putnick & Bornstein, 2016).

Measurement invariance is a necessary step to achieve before ADHD ratings can be compared adequately across ethnic groups. Although there is evidence to support that ADHD ratings may vary as a function of ethnicity, it is unclear whether these findings reflect *true* ethnic differences or are a result of non-equivalent measurement properties. In other words, groups can only be meaningfully compared when the construct being measured is the same across groups (Chen, 2008). A potential bias that may affect findings is response styles (Chen, 2008; Kline, 2016). As items in ADHD rating scales are typically rated on a Likert scale, respondents are asked to choose among a series of descriptors that best describe how much their child's behaviors correspond to the behavioral statement. For example, the teacher and parent form of the SNAP-IV has statements such as "often fidgets with hands or feet or squirms in seat," and asks parents/teachers to rate how much the statement describes their child on a 4-point Likert scale (0 = not at all, 1 = just a little, 2 = quite a bit, 3 = very much; Swanson, 1992). Cultural values may influence how individuals respond, and thus ratings may not objectively represent children's behaviors.

Studies have found that extreme response styles (ERS; i.e., choosing endpoints of a scale and avoiding midpoints) and acquiescence response styles (ARS; i.e., choosing one side of a scale consistently, which is typically on the positive end) are more frequent, or common, among Black and Hispanic/Latinx populations than other ethnic populations (Davis et al., 2011; Yang et al., 2010). On the other hand, those who identify as Asian tend to avoid ERS and ARS and select middle points on Likert scales (Davis et al., 2011; Wang et al., 2008; Yang et al., 2010). Response styles may be affected by cultural values and relatedly, social desirability and stigmatization. Hispanic/Latinx individuals may maintain ERS due to their sense of *machismo* and *la mujer,* which emphasizes male and female dominance or strength, while *familismo* may help to maintain ARS due to a reflection in social cohesion (Davis et al., 2011). Similarly, Black individuals might have higher ERS and ARS due to their desire to be accepted and perceived positively by society which is driven by their strong family values (Liu et al., 2016). As a result, Black and Hispanic/Latinx populations may be more likely to have ERS and ARS due in part to the social desirability principle—claiming attributes that are favorable while denying those that are unfavorable (Nederhof, 1985). Though the Asian group also endorses collectivistic values that are like Hispanic/Latinx and Black groups, their values in social harmony may lead them to endorse the middle group to avoid conflict and being outcasted, which may be driven by feelings of stigmatization (Wang et al., 2008).

Furthermore, another potential bias could be related to whether items are appropriately defined. More specifically, items may have different meanings across cultures (Kline, 2016). For example, in a study looking at the CBCL, Lambert and colleagues (2002) found that 25% of problems endorsed by Black parents were not included within items in the measure. Thus, items included in the scales may not encapsulate cultural definitions of what it means to be hyperactive-impulsive and inattentive (Gerdes et al., 2013; Reid et al., 1995). Additionally, the frequency of the items, or score ratings, may differ based on cultural interpretations. Depending on one's cultural norms, individuals may have different interpretations of what frequency or duration of fidgeting corresponds with "not at all", "just a little", "quite a bit", and "very much" (Reid et al., 1995). Given the earlier discussion on how ethnically minoritized parents are less likely to adopt biomedical models, they may be more tolerant of behaviors and endorse a lower frequency of behaviors (Liang et al., 2016).

Summary and Critique

ADHD rating scales are an important part of the assessment process in detecting childhood ADHD. There are both broad- and narrow-based rating scales, which either include rationally derived items for multiple psychopathologies or include DSM-derived ADHD symptom items (Collett et al., 2003; Langberg et al., 2008). In general, both broad- and narrow-band ADHD rating scales when used correctly, offer a cost-effective and feasible approach to the assessment process (Pelham et al., 2005). Advantages of using ADHD rating scales include that they are short and easy to complete, effective in detecting ADHD, and can be implemented within large populations. Additionally, although rating scales cannot be used as a replacement for an assessment to receive an ADHD diagnosis, it increases the chances of diagnosis and referral.

However, despite the clinical utility and effectiveness of using ADHD rating scales as a screening tool, there appears to be a pattern of differential ADHD ratings as a function of ethnicity. Most studies generally find that ratings are lower among Black, Hispanic/Latinx, and Asian children than White children. Extant literature does well to address the importance of attending to ethnicity in ADHD rating scales and looks at both broad- and narrow-band scales; however, several areas of improvement could help to provide further clarity in this line of research. Most studies focus primarily on Black children, with only some literature including

Asian and Hispanic/Latinx children. The limited research on Asian and Hispanic/Latinx populations makes it difficult to generalize findings. For example, although results are consistent across studies for Asian children, most, if not all, studies are not representative of the U.S. population. One study by Cuffe and colleagues (2005) attempted to use data from the NHIS, but their findings need to be interpreted with caution, given that the "other" category was not able to be analyzed as separate ethnic groups. However, as seen with how Asian and Hispanic/Latinx children have differential ratings compared with White and Black children, further inclusion of Asian and Hispanic/Latinx populations using nationally representative samples is warranted.

Additionally, studies show that the perspective of the informant, either from a teacher or parent standpoint, can differentially affect ratings. This is consistent with literature theorizing that informant discrepancies may arise from having unique perspectives on behaviors or from having different environmental expectancies (Des Los Reyes & Kazdin, 2005), as well as literature indicating that bias is a limitation for using rating scales (Pelham et al., 2005; Wexler et al., 2022). In research on parent-teacher discrepancies, parents tend to report more symptoms compared to teachers (Takeda et al., 2020). In contrast, the preliminary literature on ADHD ratings among ethnic groups points to parents rating ethnically minoritized children lower on ADHD symptoms than White children. Therefore, ethnicity may interact with parent and teacher ratings of ADHD, creating differential rating patterns than observed as a holistic unit. However, the majority of the current research tends to focus on teacher rather than parent ratings. As both parent and teacher ratings are crucial for a comprehensive examination of ADHD, there is a need for more research on ethnic differences among parent ratings of ADHD rating scales.

Given these differential ratings, the literature suggests that ADHD rating scales may maintain several weaknesses, which has multiple implications for its use across ethnic groups. A lack of measuring for functional impairment in tandem with symptomatology may limit its use across cultures, as well as undermine the importance of impairment in diagnostic decision-making (Gerdes et al., 2013; Pelham et al., 2005). Additionally, not all rating scales are accessible to ethnic groups due to language differences, and they may not be impervious to biases, making it important to include both teacher and parent reports of these scales (Harvey et al., 2013; Pelham et al., 2005).

The last weakness questions to what extent ethnicity affects the interpretation of ADHD rating scales and whether ADHD rating scales can be used equivalently across ethnic groups. Currently, there is a dearth of literature exploring this issue. A systematic review by Stevanovic and colleagues (2017) examined different child psychopathology scales, inclusive of ADHD scales such as the CBCL, SDQ, VADRS, and ADHD-RS. The authors indicated that ADHD rating scales contain mixed, weak, and inconclusive measurement invariance of these measures for use cross-culturally. Of note, many of these studies were in different countries, so there is a gap in understanding of U.S. populations (Stevanovic et al., 2017). Although some U.S. studies do indicate evidence of measurement invariance for ADHD rating scales (e.g., CRS, ADHD-RS-V, SDQ), they mostly focus on whether these measures are equivalent across White, Black, and Hispanic/Latinx populations (DuPaul et al., 2016).

In this study, I will further explore the construct validity of a broad-band rating scale, the SDQ, which has many strengths as a screening measure of ADHD: it contains an impact supplement assessing for functional impairment, is available in multiple languages, and has both a parent and teacher version of the measure. Specifically, I will focus on the factor structure of the parent report of SDQ Hyperactivity-Inattention subscale and subsequently test for measurement invariance utilizing a representative national sample that consists of White, Black, Hispanic/Latinx, and Asian ethnic groups. This will be the first step in understanding ethnic differences in ADHD symptoms since it is not conclusive whether differences in ADHD ratings across ethnic groups are a product of true cultural differences or lack of measurement invariance (Chen & West, 2008).

Strengths and Difficulties Questionnaire: Hyperactivity-Inattention Subscale (SDQ-HI)

The SDQ was developed by Robert Goodman as a broad-band rating scale used to screen and assess for a range of behaviors, emotions, and relationships among children aged 4-16 (now with versions available for 2-4, 4-17, and 18+ ages (https://www.sdqinfo.com; Goodman, 1997). Goodman based and evaluated the SDQ against the well-established Rutter questionnaires (Elander & Rutter, 1996), which was considered a reliable and valid behavioral screening measure for children (Goodman, 1997). Similar to, yet deviating from the items in the Rutter Questionnaire, the SDQ was designed to encapsulate multiple dimensions of problem behaviors, in addition to focusing on the child’s strengths, in a short and easy-to-complete format (Goodman, 1997). Currently, the SDQ is made up of 25 core items, divided into five subscales: (1) Conduct Problems, (2) Hyperactivity-Inattention, (3) Emotional Symptoms, (4) Peer Problems, and (5) Prosocial Behavior (Goodman, 1997). Along with the 25 core items, there is an additional 5-item impact supplement, which assesses for impairment (Goodman & Scott, 1999). For the purposes of this study, I will be focusing on the SDQ Hyperactivity-Inattention Subscale (SDQ-HI). This subscale is made up of five items, with two items measuring hyperactivity, two items measuring inattention, and one item measuring impulsivity (Goodman, 2001; Goodman & Scott, 1999).

Psychometric Properties of the SDQ and SDQ-HI

Reliability. Reviews have reported robust evidence that the SDQ is a psychometrically sound measure to use as a screening tool in the assessment process (Achenbach et al., 2008; Marti et al., 2022; Stone et al., 2010). In terms of reliability, the SDQ is shown to have satisfactory internal consistency, with Cronbach alpha coefficients of .82 for the total score and .77 for the hyperactivity-inattention score (Goodman, 2001). Although Goodman and colleagues used a British sample, similar Cronbach alpha coefficients were found in a U.S. sample (total score, $\alpha = .83$), helping to further provide support for satisfactory internal consistency. There is also evidence of interrater agreement between parent, teacher, and adult informants, with average Pearson correlations in clinical (r = .39-.58) and community samples (r = .24-.37; Goodman et al., 1998). Although these findings indicate low to moderate interrater agreement, they are greater than or equal to the meta-analytic mean of other child measures found by Achenbach and colleagues (1987; r = .20-.27) and by Des Los Reyes and colleagues (2015; r = .25-.30). As for test-retest reliability, the SDQ evidenced stable ratings across short and long intervals of time. Goodman (2001) found significant correlations in a 4-to-6-month period, with a mean correlation of .62 and higher correlations for total ($r = .72$) and SDQ-HI scores ($r = .72$). Further, SDQ and SDQ-HI scores remained stable over a 12-month period, where correlations for SDQ-HI ($r = .77$) and total scores ($r = .77$) remained the highest (Hawes & Dadds, 2004).

Validity. In terms of validity, the SDQ demonstrates evidence of strong convergent validity with other reliable and valid broad-band screening measures, such as the Rutter Questionnaire (total score, $r = .88$; hyperactivity-inattention score, $r = .82$; Goodman & Scott, 1999), CBCL (total score, $r = .87$; hyperactivity-inattention score, $r = .71$; Goodman, 1999, and Youth Self-report (YSR; total score, $r = .83$; hyperactivity-inattention score, $r = .76$; Kovacs &

Sharp, 2014). Therefore, the SDQ has similar content tapping multiple dimensions of children's behaviors as these screening measures. Studies using receiver operating characteristic (ROC) analyses also found that the SDQ was able to distinguish between different populations (i.e., clinical versus non-clinical), with good to excellent area under the curve coefficients (AUC; .87; .93) that are comparable to the Rutter Questionnaire (.87) and CBCL (.92), evidencing concurrent validity (Goodman, 1997; Goodman & Scott, 1999). For ADHD specifically, a study by Algorta and colleagues (2016) indicated excellent AUCs (.90-.96) across age and sex for the SDQ-HI in discriminating between children with and without ADHD. Additionally, in a U.S. sample, AUCs in the parent (.74) and child SDQ-HI (.76) were comparable to the CBCL (.81) and YSR (.77) and met the benchmark for distinguishing between ADHD and non-ADHD children (Kovacs & Sharp, 2014). Two longitudinal studies also indicated partial evidence of predictive validity, as the SDQ was shown to significantly predict a range of disorders in two-, five-, and seven-year periods, inclusive of ADHD (Croft et al., 2015; Nielsen et al., 2019).

Predictive Power. As for its predictive power, developers of the SDQ created an algorithm that combines ratings of symptoms and impairment from the SDQ across informants (i.e., parent, child, and teacher) to predict whether a child would be diagnosed with a disorder (Goodman et al., 2003). Diagnostic predictions are aggregated into three different groups: conduct and oppositional disorders, hyperactivity and inattention disorders, and anxiety, depressive, and obsessive-compulsive disorders (Goodman et al., 2003). Subscale scores of 6, 7, and 9 and impact scores of 1 or 2 are used as cutoffs to help define children into 'unlikely,' 'possible,' and 'probable' categories (see https://www.sdqinfo.com).

Using this predictive algorithm, several studies have indicated that the SDQ-HI may be useful as a screener for ADHD. In a sample made up of 10,438 children, aged 5-15, recruited from child benefit records in Britain, Goodman and colleagues (2003) found a sensitivity rate of 75% for ADHD using parent, teacher, and child reports. In another study, Ullebø and colleagues (2011) looked at a sample of 6,233 children, aged 7-9, in Norway using a longitudinal study of child mental health from 2002. The authors used the SDQ-HI and its predictive algorithm to detect ADHD in parent and teacher reports separately, with a cutoff value of 7, and reported a 59% sensitivity rate, 96% specificity rate, and .48 PPV for teacher reports and a 49% sensitivity rate, 96% specificity rate, and .44 PPV for parent-reports. Additionally, Nielsen and colleagues (2019) found a 42% sensitivity rate, 97% specificity rate, and 28.9% PPV for hyperkinetic-inattentive disorder in the ICD-10 using 1,176 parent and teacher SDQ reports from a longitudinal study in Copenhagen reassessing children at ages 5-7 and 11-12. Similar specificity and sensitivity rates can be found in other studies using samples of preschool-aged children, with varying PPVs (PPV = .31-.64; Øvergaard et al., 2018; Rimvall et al., 2014).

On the other hand, Øvergaard and colleagues (2019) analyzed the predictive power of using the SDQ-HI subscale to identify ADHD with only parent reports. Using 52,373 7-year-old children who participated in a follow-up questionnaire for the Danish National Birth Cohort, the authors found a 39% sensitivity rate, 96% specificity rate, and .09 PPV when using a cutoff score of 7. Additionally, in a U.S. sample of 6,483 children, aged 13-18, from a nationally representative study, He and colleagues (2013) found a 61% sensitivity rate, 92% specificity rate, and 0.36 PPV using only parent-reported SDQ scores and a cutoff score of 6. Despite indication that children with ADHD may be overlooked with the use of the SDQ-HI, the lower sensitivity rates and varying PPVs might be attributed with the proportion of children with ADHD in each sample (i.e., 1.8-6.7%), the type of cutoffs used, and whether multiple informants reported scores (He et al., 2013; Nielsen et al., 2019; Øvergaard et al., 2019; Ullebø et al., 2011). In consideration of these psychometric findings, and that screening is one of many steps in the assessment process to help rule out diagnoses (Mulvaney-Day et al., 2018), the SDQ-HI may be a valid screener for ADHD.

Strengths of the SDQ and SDQ-HI

Along with its psychometric properties, the SDQ has many strengths as a measure, which makes it a versatile tool used in assessment, clinical practice, and research. As benefits of the measure, the SDQ offers a way to additionally measure for impairment, provides accessibility to multiple populations, and reduces resources needed in completing assessments. Additionally, many of these benefits can help to facilitate the use of the SDQ across ethnic groups.

Impact Supplement. One of the limitations of using ADHD rating scales is that some do not offer methods of assessing for functional impairment, which results in clinicians using additional measures. In contrast, the SDQ provides a method of assessing impairment by offering an impact supplement (SDQ-I; Goodman & Scott, 1999). The SDQ-I includes five questions about the onset of difficulties, degree of distress, impairment in multiple settings (e.g., home, school, peers, leisure activities), and burden of difficulties on the family (Goodman & Scott, 1999). From the SDQ-I, an impact score can be derived by aggregating scores from the degree of distress and impairment in multiple settings (see https://www.sdqinfo.com).

Although there has not been much research regarding the validity and reliability of using the SDQ-I as a measure of functional impairment, many researchers have used this in addition to SDQ subscale scores to predict childhood disorders (Nielsen et al., 2019; Øvergaard et al., 2019; Ullebø et al., 2011). Furthermore, one study provided supportive evidence of partial validity in using the SDQ-I to assess for functional impairment. Kjærandsen and colleagues (2023) examined the Vineland Adaptive Behavior Scales II (VABS-II; Sparrow et al., 2011), Children's

Global Assessment Scale (CGAS; Shaffer et al., 1983), and SDQ impact scores among 337 children, aged 4-18, who were referred for assessment between 2012-2016 to neuropediatric clinics in Norway. The SDQ impact score correlated significantly ($p < .001$) with the VABS-II ($r = -.36$) and CGAS ($r = -.29$) impairment measures, supporting convergent validity with weak associations. Since the VABS-II and the CGAS measures different dimensions of functional impairment (i.e., adaptive functioning, general functioning) and are both clinician-rated tools, this may contribute toward the observed weak associations (Kjærandsen et al., 2023). However, although associations were weak, the SDQ-I indicated similar levels of functional impairment as the VABS-II and CGAS among children with ADHD and ASD, providing indirect support for concurrent validity (Kjærandsen et al., 2023). Additionally, albeit preliminary, the authors indicated that the SDQ-I might improve diagnostic clarity of ADHD and ASD but may not be adequate to detect diagnoses accurately when used in isolation (i.e., without the other 25 items in the scale; Kjærandsen et al., 2023). Although more research is necessary to validate the usefulness of the SDQ-I, these preliminary findings evidence a strength to using the SDQ because it offers a way to encapsulate functional impairment and improve diagnostic assessment for ADHD in one measure.

Accessibility to Multiple Populations. In addition to having an impact supplement, the SDQ is an accessible measure, increasing its utilization across a variety of contexts. The SDQ is available in three different versions; there is a parent/caregiver version, a teacher version, and a child, self-report version (Goodman, 1997; Goodman et al., 1998). Therefore, the SDQ can capture multiple perspectives of the child’s behaviors in differing environments. This helps to mitigate potential bias by providing more of a balanced view of the child’s behaviors. Furthermore, the SDQ is available to download on the official website, with over 40 languages available for use (e.g., Afrikaans, Chinese, Japanese, Spanish, Korean, Portuguese, Tagalog, Vietnamese; Goodman, 2001). Given that rating scales may not be available in certain languages, hindering its utility in populations that may not be fluent in English, the wide variety of translated versions for the SDQ addresses this limitation (Achenbach et al., 2008). Additionally, some studies provide preliminary evidence of reliability and validity, as well as conceptual and linguistic equivalence, among translated versions of the SDQ (Hassan et al., 2019; Lijuan et al., 2019; Stolk et al., 2017; Woerner et al., 2004). Although not all versions of the SDQ are well-validated and require further investigation and revision, the current work continues to facilitate the use of the SDQ among various populations.

Cost-Effectiveness. Furthermore, the SDQ is a cost-effective measure due to its affordability and brevity. The SDQ is downloadable for free via the official website (Goodman, 2001). As a result, the SDQ does not come with additional charges and helps to reduce resources needed in assessment. For example, the SDQ is a self-administered measure, meaning that training is not needed to use the measure and that the measure does not require a large quantity of staff to issue the measure. In addition, since the SDQ contains only 25 items, with only five items for the SDQ-HI, it takes less time to complete than other screening measures, such as the CBCL and ASEBA (Goodman & Scott, 1999). Also, as the SDQ-I is not a necessary component of the measure, it can be added or removed depending on the clinician's purpose for using the SDQ. Some measures of functional impairment are long, and as a result, time-consuming to include (Kjærandsen et al., 2023). However, since the SDQ-I is short to complete, it does not add excessive time that would otherwise burden the respondent. For example, longer screening measures might impede completion and affect engagement in assessment, which may make shorter screening measures more practical and feasible to implement.

Although shorter measures often sacrifice reliability and validity for conciseness, the SDQ maintained high criterion validity, when compared with a well-established semi-structured interview measurement (Parental Account of Child Symptoms; PACS; Taylor et al., 1986), than the CBCL (Goodman & Scott, 1999). In particular, the SDQ had significantly higher correlations with the PACS on the hyperactivity-inattention scales (r = .43) than the CBCL (r = .15; Goodman & Scott, 1999). Additionally, the SDQ-HI was found to have higher sensitivity than the full DSM ADHD symptom list (i.e., 68% for SDQ-HI and 54% for DSM) when ADHD was defined by a diagnostic interview measure (Posserud et al., 2014). Therefore, despite its limited number of items, the SDQ maintains its quality as a screening measure. This can be further corroborated with its other solid psychometric properties (Stone et al., 2010).

In consideration with its psychometric properties and characteristics, the SDQ is a valuable screening tool, and using the SDQ-HI can help identify childhood ADHD. Many of the strengths of the SDQ and SDQ-HI, such as its ability to measure for functional impairment, obtain perspectives from varying informants with ranging language skills, and conserve participant and clinician resources, support its acceptability as a measure for use across multiple settings and populations. This helps to reduce some of the limitations that ADHD rating scales have, which may hinder their use across differing ethnic groups. Inclusivity of impairment may not only provide greater clarity to diagnostic decisions but also caution clinicians that there may be problems despite the dismissal of symptomatology. This is especially important for ethnic minority groups that adopt less biomedical models of ADHD. Moreover, increasing access to multiple informants can reduce ethnically biased understandings toward a child's behaviors, and availability in multiple languages secures the perspectives of non-primary English-speaking parents and caregivers who wish to voice their concerns. Meanwhile, the cost-effective component of the SDQ also serves to indirectly maintain inclusion of multiple informants by allowing for administration to varying people despite time constraints or financial inability to do so.

Factor Structure of the SDQ-HI

Despite the advantages of using the SDQ to help facilitate screening of childhood disorders in diverse ethnic groups, there are some limitations to using the SDQ-HI as an ADHD rating scale. Relative to the earlier limitations, there is a lack of understanding of whether the measure is equivalent for use across ethnic groups. However, before the evaluation of measurement equivalence, the factor structure of the SDQ-HI must first be analyzed. Analyzing the factor structure of the SDQ-HI would help support conceptual understandings of ADHD and determine whether its conceptualization aligns with the DSM-V-TR to evidence the construct validity of the measure.

One-Factor SDQ-HI. No known studies have strictly analyzed the factor structure of the SDQ-HI, as studies typically group all five items in the SDQ-HI implicitly assuming a one-factor structure of the SDQ. In the original study by Goodman (2001), the factor structure of the SDQ was analyzed using exploratory factor analyses (EFAs). Results of EFAs showed that the SDQ had loaded primarily and exclusively on five predicted factors (i.e., conduct problems, hyperactivity-inattention, emotional problems, peer problems, and prosocial behaviors); each factor corresponded with one of the five subscales, with five items loaded on each subscale (Goodman, 2001). The five-factor structure of the SDQ was further replicated in studies using confirmatory factor analyses (CFAs) and exploratory structural equation modeling (ESEM; Graybill et al., 2021; Hall et al., 2019; He et al., 2013). Although some studies found evidence supporting a three-factor (i.e., externalizing, internalizing, prosocial; Dickey & Blumberg, 2004) and four-factor structure (i.e., conduct, hyperactivity, internalizing, prosocial; Bull et al., 2015), most studies generally replicate and support the five-factor structure of the SDQ (Karlsson, 2022; Marti et al., 2022; Stone et al., 2010).

As a result of these studies, the five items in the SDQ-HI are typically conceptualized as one factor, which corresponds with the unidimensional model of ADHD that was derived from the DSM-III-R (APA, 1987; Bauermeister et al., 1992). In the DSM-III-R, ADHD was not listed under different dimensions, which the DSM-III had previously defined. Instead, 14 symptoms, consisting of hyperactivity, impulsivity, and inattention, were defined under one domain, and any combination of 8 or more of these symptoms would classify a diagnosis of ADHD (Bauermeister et al., 1992). Therefore, the unidimensional model of ADHD posits that hyperactivity, impulsivity, and inattention are encapsulated in a single dimension of maladaptive behavior (Lahey et al., 1988). From a theoretical perspective, this model is supported by primary core deficit theories of ADHD. One such theory, proposed by Barkley (1997), predicates that deficiency in behavioral inhibition underlies ADHD symptomatology, as inhibition is an essential component of four executive functions (i.e., working memory, self-regulation, internalization of speech, reconstitution) that affect goal-directed behavior, task persistence, and motor control.

Two-Factor SDQ-HI. In contrast with the unidimensional model of ADHD, other researchers and clinicians view ADHD as multidimensional, with hyperactivity and impulsivity included as one factor and inattention as another factor. This two-factor structure was elicited in the DSM-IV, which categorized hyperactivity and impulsivity symptoms together while inattention symptoms remained independent from the two (APA, 1994; Bauermeister et al. 2010). This conceptualization of ADHD has stayed consistent in the DSM-IV-TR (APA, 2000), DSM-V (APA, 2013), and the DSM-V-TR (APA, 2022). Further, predominant presentations of the disorder were derived from the separation of these dimensions (i.e., combined, predominantly hyperactive, predominantly inattentive; APA, 2022).

Changes in the DSM's conceptualization of ADHD were prompted by factor analyses of the ADHD symptom list suggesting two-factor solutions (Bauermeister et al., 1992; Lahey et al., 1988; Pillow et al., 1997). Additionally, initial field trials for the DSM-IV provided evidence that hyperactivity-impulsivity and inattention were separate dimensions, as they were associated with different types of impairment, and clinicians diagnosed children with ADHD independent of inattention symptoms (Lahey et al., 1994). In support of a two-factor structure of ADHD, genetic studies analyzing etiological influences of ADHD indicate that there are both shared and unique genetic relationships among hyperactivity-impulsivity and inattention domains (Kuntsi et al., 2015; Larsson et al., 2006). Further, studies show that inattention is associated with executive dysfunction while hyperactivity-impulsivity is associated with temporal discounting, or delay aversion (Chhabildas et al., 2001; Scheres et al., 2008; Sonuga-Barke et al., 2003). This relates to the dual-pathway theory of ADHD that supposes that differing brain circuits (i.e., meso-cortical, meso-limbic) form separate executive and motivational pathways that both give rise to the development of heterogeneous ADHD symptomatology (Sonuga-Barke, 2002). In other words, these studies theorize that hyperactivity-impulsivity and inattention are overlapping, yet distinct entities due to their differing etiological pathways.

In examining the factor structure of varying measures of ADHD, some studies support a unidimensional conceptualization of ADHD (Bauermeister, 1992; Flory et al., 2020); however, results from these studies may have been influenced by the age group of their samples. In other words, since the studies utilized preschool-aged and college-aged individuals, there may have been a lack of heterogeneity in symptomatology, making a unidimensional structure more suitable. This coincides with developmental trajectories of ADHD describing how hyperactive-impulsive symptoms are predominantly present in preschool-aged children and decline with age, along with inattention symptoms becoming more present over time (Biederman, 2000; Cherkasova et al., 2013). Therefore, multidimensional conceptualizations of ADHD may be more appropriate for a 6- to 13-year-old age range (Bauermeister, 1992). Furthermore, most other studies examining the factor structure of ADHD favor the two-factor model (Arias et al., 2016; Bauermeister et al., 1992; Bauermeister et al., 2010). In a systematic review, the two-factor model of ADHD was supported cross-culturally in at least 15 countries using a range of informants and rating scales (Bauermeister et al., 2010). As such, ADHD is conceptualized best as a multidimensional, two-factor construct, and it may be appropriate to examine whether the SDQ-HI best fits a one or two-factor structure as well.

Measurement Invariance of the SDQ-HI

After exploring the factor structure of the SDQ-HI by comparing one- and two-factor models, other necessary steps are needed to qualify the use of this measure across a variety of people and contexts. Previous research has indicated that the five-factor structure of the SDQ is invariant across age, sex, treatment groups, time, and informants (i.e., mothers and fathers, parents and teachers; Chiorri et al., 2015; Graybill et al., 2021; Hall et al., 2019; He et al., 2013; Rogge et al., 2018; Staatz et al., 2023). However, less research has attempted to understand whether the constructs being captured within the SDQ in general, and those of the SDQ-HI in particular, are equivalent across ethnic groups.

In the U.S., two studies have examined the measurement invariance of the SDQ with ethnicity in mind. In a study by He and colleagues (2013) that looked at a sample of 6,483 13–

18-year-old children from the National Comorbidity Survey – Adolescent Supplement, they found evidence of partial measurement invariance (i.e., configural, metric, scalar invariance) for the five-factor, parent-reported SDQ across White, Black, and Hispanic children. Although there were other ethnic groups available for analysis, they were not included due to the difficulty in parsing each group from each other (He et al., 2013). In a different study by Graybill and colleagues (2021), who examined 1,545 middle-school children from a federally funded project implementing universal health screening in 2015 and 2016, they also supported partial measurement invariance for the five-factor, parent-reported SDQ across White and Black children. In a third study, Gomez and Stavropoulos (2018) attempted to assess measurement invariance in the parent-reported SDQ (both in English and Malay versions) among a sample of 1,407 children, aged 5-13, in Malaysia. Results from their study indicated full measurement invariance (i.e., configural, metric, scale, residual invariance) for the parent-reported SDQ across Malay, Chinese, and Indian Malaysian children. However, since their study utilized a two-factor structure (i.e., positive construal factor and psychopathology factor), which separated SDQ-HI items across factors and removed two non-salient SDQ-HI items, this does not provide a clear understanding of whether the SDQ-HI is invariant across ethnicity in the U.S.

Summary and Critique

Many studies have supported the validity and reliability of the SDQ, with studies also evidencing predictive power in using the SDQ-HI to screen for ADHD. In addition, with its strong psychometric properties, the SDQ offers many benefits over other rating scales: it includes items measuring for impact as a supplemental measure, is available in multiple languages and versions, and is cost-effective. However, despite its many strengths, the SDQ, and particularly the SDQ-HI requires further psychometric evaluation to be able to be adequately used as an ADHD rating scale among diverse ethnic populations. More specifically, there is a need to assess both the factor structure of the SDQ-HI and whether its measurement properties are invariant across ethnicity.

To my knowledge, the factor structure of the SDQ-HI has yet to be explored in prior studies. Given that the SDQ-HI is typically seen as one factor within the broader, five-factor SDQ model, and that there are only five items in the scale, the SDQ-HI may maintain a one-factor structure. However, research has yet to compare one- and two-factor models of the SDQ-HI, so it is unclear whether the ADHD construct conceptualized in the scale is unidimensional or two-dimensional. Although studies that analyzed the SDQ structure show that the five items load strongly on one factor, there is some evidence to support that the SDQ-HI may fit better with a two-factor structure. Given the developmental course of ADHD discussed earlier and the age range that the measure is used for (Goodman, 1997), the SDQ-HI might be more consistent with a two-factor structure. Further, when developers of the SDQ were creating items for the SDQ-HI, they based the items off symptom domains from the DSM-IV, which conceptualized ADHD as multidimensional: two items encapsulated inattention and three encapsulated hyperactivity-impulsivity (Goodman, 2001). Additionally, some studies have provided evidence of minor factors in the SDQ-HI with added modifications of covariances between two items measuring inattention and the two items measuring hyperactivity helping to improve the model fit of the measure (Bøe et al., 2016; Karlsson et al., 2022). Since no known studies have investigated the two-factor structure of the SDQ-HI, more research comparing the one- versus two-factor structure is needed. Doing so will help to implicate whether the SDQ-HI aligns well with the current conceptualization of ADHD in the DSM and provide greater support toward using it as a screening tool.

No studies have investigated whether the measurement properties of SDQ-HI are invariant across ethnicity. Although previous studies help to evidence that the five-factor model of SDQ is invariant across ethnicity, they focused on White, Black, and Hispanic/Latinx child populations, overlooking other ethnic groups. Further, the existing study including Asian populations in other countries may not be generalizable to the U.S. context. To date, no other studies have attempted to analyze measurement invariance of the SDQ-HI across ethnicity in a U.S. population, making it difficult to make inferences about whether the measure is invariant for Asian children. Additionally, the current literature conceptualized the SDQ-HI as a single-factor model within the broad five-factor structure of the SDQ, except for Gomez and Stavropoulos (2018); therefore, results may differ if the SDQ-HI is analyzed separately and fits a two-factor structure instead. In consideration of these limitations, more work is needed to support whether the SDQ-HI can be used appropriately across ethnic groups without introducing biases.

Aims and Hypotheses

My current study had three primary aims: (1) identifying the factor structure of the SDQ-HI, (2) analyzing measurement invariance of the SDQ-HI across White, Black, Hispanic/Latinx, and Asian children, and (3) examining latent mean differences of the SDQ-HI across the four ethnic groups.

Aim 1: Factor Structure of the SDQ-HI

In the literature, there is debate about whether ADHD should be conceptualized as unidimensional or multidimensional; however, studies generally support a two-factor structure, coinciding with recent versions of the DSM-V-TR (Bauermeister et al., 2010). Given that there are no known studies analyzing the specific factor structure of the SDQ-HI, it is unclear whether a one- or two- factor structure fits this subscale best. Therefore, the first aim of the study was to investigate the factor structure of the SDQ-HI by comparing the fit of one- versus two-factor models. Since the SDQ-HI is used with an age range that typically displays heterogenous symptoms across hyperactivity-impulsivity and inattention domains (Bauermeister, 1992; Goodman, 1997), was designed with a multidimensional conceptualization of ADHD (i.e., having items correspond with hyperactive, impulsive, and inattentive symptom domains from the DSM-IV; Goodman, 2001), and evidenced possible subfactors (Bøe et al., 2016; Karlsson et al., 2022), I predicted that a two-factor model of the SDQ-HI would fit the data better than a one-factor model.

Aim 2: Measurement Invariance of the SDQ-HI

Although studies indicated that there may be potential cultural factors affecting how ethnic groups interpret and respond to measures (Davis et al., 2011; Yang et al., 2010), research on measurement invariance of the SDQ-HI is sparse. The second aim of the study was to examine whether the identified factor structure of the SDQ-HI in the first step would hold invariant properties across the four ethnic groups. No hypothesis was proposed due to the limited evidence.

Aim 3: Latent Mean Differences across SDQ-HI Scores

Studies on ethnic differences in ADHD rating scales find that Black, Hispanic/Latinx, and Asian children were rated lower than White children, specifically when rated by their parents (Cuffe et al., 2005; DuPaul et al., 2016). However, since there is a lack of studies analyzing whether these rating scales are equivalent across ethnicity (Stevanovic et al., 2017), it is unclear whether these results reflect *true* differences or biases in measurement. Assuming the SDQ-HI could achieve at least configural invariance in the analyses mentioned above, this study investigated latent group mean differences across ethnicity. In consideration of stigmatized perceptions of ADHD across ethnic minority groups, as well as adherence to less biomedical models of ADHD, having less education about the disorder, and higher distress thresholds (Liang et al., 2016), it was hypothesized that Black, Hispanic/Latinx, and Asian children will have lower latent means than White children.

Methods

Data and Sample

The data for this study was derived from the Sample Child portion of the 2019 National Health Interview Survey (NHIS), a nationally representative cross-sectional household interview survey conducted annually by the National Center for Health Statistics (NCHS) that includes data on a broad range of health and socio-demographic information. The NCHS uses geographically clustered, multistage sampling techniques to obtain representative samples of the civilian, noninstitutionalized U.S. population. The U.S. is first partitioned into 1,689 geographic areas—counties or equivalents that are contiguous and do not cross state borders—which either make up one stratum or are divided into two strata dependent upon the population density. Each stratum is further divided into clusters that contain approximately 2,500 addresses, and these clusters are systematically selected to comprise the sample. For the 2019 NHIS, 534 clusters of addresses within 310 geographic areas were selected, and these were inclusive of each state. After collecting basic demographic information about everyone typically residing in each household address, one child (i.e., aged 17 years or younger) is randomly selected, and a knowledgeable adult or parent (from here on referred to as parent) is asked to provide information about that child in a face-to-face or telephone interview. Their responses to these questions make up the data in the Sample Child portion of the NHIS.

The 2019 NHIS originally included 9,193 children from 33,138 households. Children aged 0-3 (*n* = 1,985) children are excluded from this study since the Strengths and Difficulties Questionnaire (SDQ) measure only contains information about children aged 4-17. Additionally, due to the small amount of data available for American Indian/Alaska Native children (*n* = 53) and the inability to disaggregate ethnicity from the other category (*n* = 374), these two groups are excluded from analyses. Further, cases that were missing values on items in the SDQ-HI (*n* = 108) were deleted from analyses. With these exclusions, the total analytic sample is 6,673 children (Non-Hispanic White: *n* = 3,779; Hispanic/Latinx: *n* = 1,670; Non-Hispanic Black: *n* = 828; Non-Hispanic Asian: *n* = 396).

Measures

The primary measure is the Strengths and Difficulties Questionnaire (SDQ; see Appendix A), a measure developed by Robert Goodman that assesses a child's emotional and behavioral functioning (Goodman, 1997). This 25-item measure is used as a screening tool for clinical assessment and is composed of five subscales looking at conduct problems, hyperactivity-inattention, emotional symptoms, peer problems, and prosocial behavior. Prior studies demonstrate that the SDQ is a psychometrically solid measure, with evidence of both validity and reliability (see *Psychometric Properties of the SDQ*). For the current study, I will specifically be looking at the five items comprising the SDQ-HI. Table 1 presents the SDQ-HI items. There are similar versions of the measure used by parents, teachers, and children; however, the 2019 NHIS only provides the parent-report version of the SDQ. In this version, each parent, or knowledgeable adult, was asked to rate how well each statement matches their child's behavior in the past six months on a 3-point Likert scale (1 = not true, 2 = somewhat true, 3 = certainly true). Additionally, some items were reverse-coded for positively worded items. The scores from each item were then added together to provide a total SDQ-HI subscale score, with a scale between 0-10. A cutoff score of 7 is typically used in epidemiological studies to classify whether one may qualify for an ADHD diagnosis (Overgaard et al., 2019; Ullebo et al., 2011).

The grouping variable in this study was a combined variable of the child's race and Hispanic/Latinx ethnicity. The child's race and Hispanic/Latinx ethnicity were reported by a parent or knowledgeable adult during the interview. The variable was coded as follows: 1 = Non-Hispanic White (i.e., White group), 2 = Hispanic/Latinx, 3 = Non-Hispanic Black (i.e., Black group), and 4 = Non-Hispanic Asian (i.e., Asian group).

In addition, other demographic and socioeconomic variables were included about the child: sex (1 = male, 2 = female), age (4-17), nativity status (0 = U.S. born, 1 = foreign-born, -9 = missing), the highest level of education attained by an adult in the family's household (1 = grade 1-12, no diploma, 2 = high school graduate or GED, 3 = some college, no degree, 4 = associate's degree, 5 = bachelor's degree, 6 = master's degree, 7 = professional school or doctoral degree, -9 = missing), total combined family income (1 = 0-34,999, 2 = 25,000-49,999, 3 = 50,000-74,999, 4 = 75,000-99,000, 5 = 100,000 and over, and -9 = missing), and previous ADHD diagnosis (0 = no, 1 = yes, -9 = missing) are accounted for. Additionally, the child's relationship to the adult informant was reported (1 = parent [biological, adoptive, or step], 2 = grandparent, 3 = aunt/uncle, 4 = brother/sister, 5 = other relative/non-relative, -9 = missing).

Analytic Approach

Preliminary Analyses

Exploratory Analyses

Distributional properties, such as outliers, normality, skewness, and kurtosis were examined for each study variable.

Descriptive Statistics

Demographic and socioeconomic information, including sex, age, nativity status, highest household education, family income, and ADHD diagnosis were analyzed to describe the sample's characteristics. Further, minimum, maximum, mean, and standard deviation values for the SDQ-HI were examined.

Aim 1: Factor Structure of the SDQ-HI

A series of confirmatory factor analyses (CFAs) was used to identify whether the SDQ-HI best fits a one- or two-factor structure. Analyses were performed via MPlus (version 8.9; Muthén & Múthen, 1998-2023). Since the items in the SDQ-HI are rated on a 3-point Likert scale, the variables were treated as ordinal, and weighted least squares estimation (WLSMV) was employed. Since the NHIS uses a complex, multistage probability sample, NCHS (2020) recommends that researchers account for stratification, clustering, and sampling weights in their analyses. As such, I included the corresponding variables provided in the data and indicated the type of analysis as 'complex.' χ^2 test of model fit is typically used to determine model fit, where statistical significance suggests that the model does not provide a perfect fit to the observed data (Mueller, 2000). However, since χ^2 is sensitive toward large sample sizes, this was not heavily relied upon. Instead, root mean square error of approximation (RMSEA), standardized root mean squared residual (SRMR), comparative fit index (CFI), Tucker-Lewis index (TLI) were used to determine model fit. Additionally, Hu and Bentler (1999) recommended that good, acceptable fit is indicated by an RMSEA value of .06 or lower, an SRMR value lower than .08, and CFI and TLI values that are .95 or higher. Given that the models are nested WLSMV was employed, χ^2 difference tests were conducted to compare model fits. A *p*-level value of .05 was used to determine statistical significance, where the model with more parameters is supported.

Aim 2: Measurement Invariance of the SDQ-HI

Upon identifying the best fitting factor structure (i.e., baseline model) for each ethnic group, a series of CFAs were used to test for configural invariance by freeing all factor loadings and item thresholds. Model fit indices were compared with the acceptable indices proposed by Hu and Bentler (1999; RMSEA ≤ .06, SRMR < .08, CFI ≥ .95, TLI ≥ .95). Upon determining whether configural invariance was met, the alignment optimization approach was employed using a fixed specification and Whites as the reference group (i.e., means and variances of the reference group fixed to 0 and 1, respectively; Asparouhov & Muthén, 2023). Approximate measurement invariance was determined by examining the total number of non-invariant item parameters and their degree of non-invariance (e.g., R^2). According to Asparouhov & Muthén (2023), higher R^2 values closer to 1 typically indicate invariant parameters, while lower R^2 values closer to 0 indicate non-invariant parameters. However, there are several limitations to using R^2 as a measure for invariance, such as having low values despite parameter estimates being invariant among groups and values of 0 due to large variance (Asparouhov & Muthén, 2023). Therefore, pairwise comparisons of factor loadings and item thresholds were examined along with the more restrictive rule of thumb suggested by Muthén & Asparouhov (2014), where no more than 25% of items should be non-invariant.

Aim 3: Latent Mean Differences across SDQ-HI Scores

After establishing whether the alignment optimization approach performed well and that the factor means were able to be compared meaningfully, pairwise comparisons of latent means were examined. A *p*-value of .05 was used to determine statistical significance.

Results

Preliminary Analyses

Exploratory Analyses

The results of exploratory analyses showed that the statistics related to normality, outliers, skewness, and kurtosis, were all within the acceptable range or below the recommended cut-offs.

Descriptive Statistics

Descriptive statistics are reported in Table 2. Children in the sample were aged 4-17 (*M* = 11.1 years, *SD* = 4.06). The sample was 56.4% White, followed by 25.2% Hispanic/Latinx, 12.4% Black, and 6% Asian. Children were predominantly male (51.9%) and U.S. born (94.8%). The adult informant was primarily the child's parent (i.e., biological, adoptive, or step; 93.5%), followed by a grandparent (4.2%), brother/sister (0.8%), other relative/non-relative (0.8%), and aunt/uncle (0.6%). According to the adult informant, the highest level of education in their family was a bachelor's degree (25.3%) followed by high school graduate or GED (17.3%), master's degree (16%), associate's degree (15.9%), some college (14.4%), professional or doctoral degree (5.7%), and grades 1-12 (5.2%). Additionally, the sample had a total combined family income between 1000,000 and over (37.5%), 0-34,999 (20.7%), 50,000-74,999 (17.1%), 75,000-99,999 (13.1%), and 35,000-49,999 (11.5%). Furthermore, 10.3% of children in the sample were previously diagnosed with ADHD.

Aim 1: SDQ-HI Factor Structure

SDQ-HI factor structure for the entire sample. The first set of exploratory analyses investigated the factor structure for SDQ-HI for the entire sample. Table 3 presents model fit indices for the one-factor (Model 1; see Figure 1) and two-factor models (Model 2; see Figure 2) of the SDQ-HI. The fit indices indicated mediocre fit for Model 1 (RMSEA = .115, SRMR = .045, CFI = .967, TLI = .934) and Model 2 (RMSEA = .118, SRMR = .042, CFI = .972, TLI = .930). However, the χ^2 difference test suggested that Model 2 had a significantly better model fit compared to Model 1 [$\Delta\chi^2(1) = 79.767$, $p < .001$].

To improve model fit, modification indices (MIs) of correlated residual terms were examined. The correlated residual terms for SDQ1 (i.e., *good attention*) and SDQ5 (i.e., *thinks first*) had large MIs across models (MI = 319.844; MI = 188.336). Given that these two items are the only reverse-coded items in the SDQ-HI, both models were modified to include the correlated residual term between these two items (see Figures 3 and 4). Model fit indices indicated mediocre fit for the modified one-factor model (see Table 3; Model 3; RMSEA = .104, SRMR = .032, CFI = .978, TLI = .946) but reasonable fit for the modified two-factor model (Model 4; RMSEA = .058, SRMR = .017, CFI = .997, TLI = .983). Further, the data fit the modified models significantly better [$\Delta\chi^2(1) = 227.266$, $p < .001$; $\Delta\chi^2(1) = 142.341$, $p < .001$]. When Model 3 and Model 4 were compared, the Model 4 had a significantly better fit [$\Delta\chi^2(1) = 184.651$, $p < .001$]. Overall, this suggested that the SDQ-HI fits a two-factor structure better[2].

Testing the SDQ-HI factor structure for each ethnic group. A series of CFAs were conducted to determine the best factor structure for each ethnic group, respectively (see Table 4). Given the results based on the total sample, the modified versions of the one- and two-factor models (i.e., Model 3 and Model 4 with correlated residuals) were compared. Overall, Model 3 provided mediocre fit for each group (White: RMSEA = .140, SRMR = .032, CFI = .982, TLI = .956; Hispanic/Latinx: RMSEA = .115, SRMR = .035, CFI = .974, TLI = 0.936; Black: RMSEA = .140, SRMR = .035, CFI = .974, TLI = .936; Asian: RMSEA = .085, SRMR = .043, CFI =

[2] Although results suggest that a two-factor structure fits best, the inattention and hyperactivity-impulsivity factors are highly correlated with one another; thus, caution should be taken when interpreting results.

.979, TLI = .947). In contrast, Model 4 had reasonable fit for each group (White: RMSEA = .087, SRMR = .019, CFI = .995, TLI = .983; Hispanic/Latinx: RMSEA = .067, SRMR = .019, CFI = .993, TLI = .978; Black: RMSEA = .076, SRMR = .020, CFI = .995, TLI = .982; Asian: RMSEA = .065, SRMR = .030, CFI = .991, TLI = .969). Additionally, Model 4 fit the data significantly better than Model 3 for White [$\Delta\chi^2(1) = 180.663$, $p < .001$], Hispanic/Latinx [$\Delta\chi^2(1) = 61.060$, $p < .001$], Black [$\Delta\chi^2(1) = 41.272$, $p < .001$], and Asian groups [$\Delta\chi^2(1) = 6.574$, $p < .05$]. Therefore, Model 4 was retained as the baseline model for further analyses, and results indicated that the two-factor structure is consistent across ethnic groups.

Aim 2: Measurement Invariance of the SDQ-HI

Testing the configural invariance model. To test whether the alignment optimization approach could be used, all the loadings and thresholds were freed in the baseline configural invariance model, with all ethnic groups included. The configural invariance model has reasonable fit (see Table 5; RMSEA = .078, SRMR = .020, CFI = .995, TLI = .983), forming the basis for alignment optimization. This also suggested that the SDQ-HI captures both inattention and hyperactivity-impulsivity across ethnicity.

Alignment optimization. The alignment optimization approach was used with a fixed specification, and the White group was used as a reference whose factor mean was restricted to 0. Table 6 presents standardized item loadings and item thresholds for each ethnic group. The results demonstrated invariance for all item loadings across ethnic groups (*ps* > .05). Additionally, R^2 values for tests of item loadings ranged from .521 to .999, except for that of SDQ5 (0). These results demonstrated strong support for metric invariance of the SDQ-HI such that ethnic groups had similar meanings of inattention and hyperactivity-impulsivity.

With regard to the test of invariance for item thresholds, the thresholds for SDQ3, SDQ4, and SDQ5 significantly varied across ethnic groups (*ps* < .05). Indicated in bold, the second threshold of SDQ3 (i.e., SDQ3$2) was non-invariant for the Asian group, while the second threshold of SDQ4 (i.e., SDQ4$2) and the first threshold of SDQ5 (i.e., SDQ5$1) was non-invariant for the Hispanic/Latinx group. Pairwise comparisons indicated that the second threshold of SDQ3 significantly differed in the Asian versus Black comparison ($\Delta\tau = -1.427$, $p < .05$), Asian versus White comparison ($\Delta\tau = -1.174$, $p < .001$), and Hispanic/Latinx versus White comparison ($\Delta\tau = -.852$, $p < .01$). Further, the second threshold of SDQ4 significantly differed in the Asian versus White comparison ($\Delta\tau = -.523$, $p < .05$), Hispanic/Latinx versus White comparison ($\Delta\tau = -.487$, p < .001), and Black versus White comparison ($\Delta\tau = -.386$, $p < .05$). Lastly, the first threshold of SDQ5 significantly differed in the Black versus Hispanic/Latinx comparison ($\Delta\tau = -.308$, $p < .01$), Asian versus White comparison ($\Delta\tau = .259$, $p < .05$), and Hispanic/Latinx versus White comparison ($\Delta\tau = .249$, $p < .01$). Given non-invariance in thresholds for three of the five items (3/5 = 60% non-invariance), approximate scalar invariance was not met. In general, Asian children were significantly more likely to be perceived as fidgety (i.e., SDQ3) than White and Black children despite displaying similar behaviors. Additionally, Hispanic/Latinx children were more likely to be perceived as restless (i.e., SDQ4) than White children and less likely to be perceived as impulsive (i.e., SDQ5) than White and Black children despite displaying similar behaviors.

Aim 3: Latent Mean Differences Across Ethnicity

Although the alignment optimization indicated there might be potential bias in item thresholds across the four ethnic groups, we continued to analyze for latent mean differences. To examine latent mean differences, additional pairwise analyses were conducted between the

White reference group and Hispanic/Latinx, Black, and Asian groups separately[3]. Table 7 includes results for the model fit assessment of the configural invariance models; Table 8 presented results for the item loadings and item thresholds; and Table 9 reports latent group mean comparisons.

White vs. Hispanic/Latinx. In the White versus Hispanic/Latinx model, configural invariance was achieved (RMSEA = .082, SRMR = .019, CFI = .995, TLI = 0.982). When using the alignment optimization, none of the item loadings significantly differed between the two groups. Additionally, the R^2 values were relatively high (.849 to 1) except for SDQ5 (0). When examining the item thresholds, the first threshold of SDQ1 ($\Delta\tau$ = .244, $p < .001$), the second threshold of SDQ4 ($\Delta\tau$ = -.430, $p < .001$), and the first threshold of SDQ5 ($\Delta\tau$ = .275, $p < .01$) varied significantly. The 25% cut-off (3/5 = 20% non-invariance) was not met, which suggested that there was a lack of scalar invariance between the two groups. Additionally, there was no significant difference in the latent means for inattention ($M_{Hispanic}$ = .093, $p > .05$) and hyperactivity-impulsivity ($M_{Hispanic}$ = .062, $p > .05$), when items with non-invariant parameters were excluded.

White vs. Black. In the White versus Black model, configural invariance was achieved (RMSEA = .084, SRMR = .019, CFI = .995, TLI = 0.984). The alignment optimization indicated that none of the item loadings and thresholds significantly differed, proving support for metric and scalar invariance. Comparisons of latent means were nonsignificant for inattention (M_{Black} = .127, $p > .05$) and hyperactivity-impulsivity (M_{Black} = -.055, $p > .05$).

White vs. Asian. In the White versus Asian model, configural invariance was achieved (RMSEA = .077, SRMR = .020, CFI = .996, TLI = .987). The results from the alignment

[3] Latent mean comparisons across all ethnic groups are reported in the Supplementary Material.

optimization showed that all factor loadings did not significantly differ. When comparing the item thresholds, there was a significant difference found for SDQ3$2 ($\Delta\tau = -1.104$, $p < .01$), but the amount of non-invariance nevertheless met the 25% cut-off (1/5 = 20% non-invariance). Results from latent mean comparisons indicated no significant differences in inattention ($M_{Asian} = -.347$, $p > .05$) and hyperactivity scores ($M_{Asian} = -.228$, $p > .05$), when items with non-invariant parameters were excluded.

In sum, across all pairwise comparisons, there does not appear to be significant differences in inattention. Additionally, in the White vs. Black comparison, there were no significant differences in hyperactivity-impulsivity. However, in the White vs. Hispanic/Latinx comparison and White vs. Asian comparison, there was a lack of scalar invariance for hyperactivity-impulsivity making it difficult to interpret latent mean difference results in this domain.

Discussion

The current study aimed to further understand the construct validity of the SDQ-HI by first examining whether the factor structure best fits a one- or two-factor model using a nationally representative community sample of children. Additionally, this study aimed to assess the measurement invariance of the SDQ-HI in White, Black, Hispanic/Latinx, and Asian children. After investigating the invariance of the SDQ-HI across ethnic groups, this study further analyzed whether latent group means differed across ethnicities. Consistent with the first hypothesis, results suggested that the two-factor model fit the SDQ-HI better than the one-factor model. This was consistent across White, Black, Hispanic/Latinx, and Asian ethnic groups. Additionally, upon examining configural invariance and using the alignment optimization approach, results indicated that overall, the factor structure and item loadings of the SDQ-HI were relatively invariant across all ethnic groups. However, as some item thresholds were found to be non-invariant, this suggested that there were differences in the way that parents perceived children's behaviors across ethnicities, particularly for items capturing hyperactivity-impulsivity. Thus, although it was hypothesized that Black, Hispanic/Latinx, and Asian children would have lower latent means than White children, the results suggested that latent means for inattention and hyperactivity did not significantly vary between White and ethnically minoritized groups. It is possible that ethnic biases related to perceptions in behaviors could have affected these results.

Factor Structure and Contributions to ADHD's Conceptualization

To our knowledge, this is the first study to analyze the factor structure of the SDQ-HI, given that previous studies have primarily examined the factor structure of the SDQ in its entirety (Graybill et al., 2021; Hall et al., 2019; Karlsson, 2022; Marti et al., 2022). Although the SDQ-HI was developed with a unidimensional model of ADHD in mind, the current study's findings are in line with research suggesting that ADHD may be better conceptualized as a multidimensional construct, with inattention and hyperactivity-impulsivity as two separate but related factors (Bauermeister et al., 2010; Sonuga-Barke, 2002). More specifically, prior studies have theorized that there may be differing neural pathways to the disorder and unique genetic relationships within each symptom domain (Kuntsi et al., 2015; Sonuga-Barke, 2002).

Alongside these broad understandings of the factor structure, the study findings indicate that the two-factor model best fits the SDQ-HI compared to the one-factor model across ethnicity. This is related to prior work indicating stability in the two-factor structure of ADHD across different countries/cultural groups using other ADHD rating scales (Bauermeister et al., 2010; Gómez-Benito et al., 2015). A multidimensional, two-factor structure of ADHD may better suit various ethnic groups as it helps to encapsulate alternative presentations of ADHD that may be potentially influenced by culture. However, to understand whether the same factors and their associated items are truly meaningful to each group, we needed to determine whether the SDQ-HI was invariant across ethnicity.

Equivalence of SDQ-HI across Ethnicity

Consistent with prior work suggesting that the SDQ is invariant across ethnic groups (Graybill et al., 2021; Gomez & Stravropoulos, 2018; He et al., 2013), the results from this study evidenced that the SDQ-HI is invariant across White, Black, Hispanic/Latinx, and Asian children. The current work adds to this body of literature as it examined whether the two-factor structure of the SDQ-HI, in particular, was equivalent and used a wider range of ethnic groups. As findings indicated that the SDQ-HI met configural invariance and alignment optimization results provided additional support for factor loading invariance (i.e., metric invariance), this suggests that the SDQ-HI measures the same constructs (i.e., inattention, hyperactivity-impulsivity) and has similar relevance or meaning across each ethnic group. Moreover, the findings demonstrate that there is a commonality in symptom meanings and presentations across inattentive and hyperactive-impulsive domains. However, it is important to note that this does not underestimate the role that culture may play in ADHD. Culture may still impact the expression of ADHD through risk markers and protective factors, as previously evidenced in research on other mental health conditions (Canino et al., 2008; Patil et al., 2018). Additionally, culture may affect interpretations of what is considered normative versus deviant behavior or the ways that people typically respond to measures in general, which in turn might be associated with the utilization of mental health services.

Influence of Culture and Ethnicity on Behavior Endorsements

Although the factor structure and factor loadings of the SDQ-HI were invariant across ethnicity, the results showed that some of the item thresholds were non-invariant (i.e., lack of scalar invariance), particularly those related to hyperactivity-impulsivity. In other words, there were significant differences in how ethnic groups perceived the severity of these behaviors. Asian children were more likely to be perceived as fidgety than White children and Hispanic/Latinx children were more likely to be perceived as restless than White and Black children. Conversely, Hispanic/Latinx children were less likely to be perceived as impulsive compared to White and Black children. These findings align with prior research suggesting differing perceptions of behaviors within the hyperactivity-impulsivity domain among Hispanic/Latinx and Asian groups (Arcia & Fernandez, 2003; Gerdes et al., 2013; Watabe et al., 2022).

Multiple interpretations may help us understand why these item thresholds varied. First, there may be different understandings of what people consider as “not true,” “somewhat true,” and “certainly true”—response choices that are provided for each item—based on the child’s ethnic background. Previous work suggests that people may have different tolerance levels or perspectives on what is deemed appropriate or inappropriate based on ethnicity (Liang et al., 2016). Item scales for Asian and Hispanic/Latinx children on hyperactive behaviors may be lower than other ethnic groups (e.g., White and Black children) due to deviance from cultural norms and disciplinary practices. In Asian and Hispanic/Latinx cultures, social cohesion and social harmony are highly valued due to collectivistic mindsets (Davis et al., 2011; Wang et al., 2008). Therefore, hyperactive-impulsive behaviors that deviate from these cultural values may be less tolerable by these ethnic groups.

On the other hand, the finding that Hispanic/Latinx children may be less likely to be perceived as impulsive aligns with prior work indicating that Hispanic/Latinx parents may not label or endorse their children as “impulsive” despite displaying these behaviors (Arcia & Fernandez, 2003; Haack et al., 2018). Studies have suggested that Hispanic/Latinx parents may not endorse items unless they perceive it to impact important cultural values, such as *familismo* and *respecto* (Haack et al., 2018). Therefore, they might recognize impairment more than certain symptoms. Alternatively, response styles may affect item endorsements. Previous studies have found that Hispanic/Latinx individuals tend to respond to measures by choosing endpoints of a scale (i.e., extreme response styles), which might be related to social desirability and wanting to assert *machismo* and *la mujer* (i.e., male and female dominance or strength; Davis et al., 2011; Yang et al., 2010). Thus, response styles may help to account for the differences seen in the thresholds across items for Hispanic/Latinx children due to tendencies to respond to extreme ends of the scale.

The differences in thresholds highlight the need to still consider ethnicity and culture when using the SDQ-HI. However, in general, results still support the use of the SDQ-HI as a screening tool for ADHD, which corresponds with other research highlighting strong reliability and validity of scores produced by the measure (Achenbach et al., 2008; Algorta et al., 2016; Kovacs & Sharp, 2014; Marti et al., 2022; Øvergaard et al., 2018; Stone et al., 2010; Ullebø et al., 2011). Our study adds another component to this line of research by increasing the construct validity of the SDQ-HI across different ethnic groups.

Comparing Latent Means of Inattention and Hyperactivity Across Ethnicity

Despite results suggesting a lack of scalar invariance, we decided to continue looking at pairwise comparisons of latent mean differences in inattention and hyperactivity-impulsivity.

These findings add to the body of work on ethnic differences in ADHD ratings, as it expands our understanding of other ethnic groups (i.e., Hispanic/Latinx, Asian) and includes a different informant perspective than teachers (i.e., parent or knowledgeable adult). In contrast with studies suggesting that Asian, Hispanic/Latinx, and Black children have lower parent ADHD ratings compared to White children (Cuffe et al., 2005; DuPaul et al., 2016; Morgan et al., 2016; Wexler et al., 2022), there were no differences in latent means of inattention between ethnic groups found in the current study. The findings also indicate that Black and White children, displaying both inattentive and hyperactive-impulsive behaviors, may have similar informant ratings when the constructs in the measures are meaningful and relevant to them. This could be said given that the SDQ-HI was found to be equivalent across both groups. However, the same could not be interpreted for Hispanic/Latinx and Asian children compared to White children, given that there was evidence of non-invariant item thresholds for the hyperactivity-impulsivity domain described earlier. As the alignment optimization does not account for non-invariant items when computing latent means and there are only five items in the measure, latent means may not capture the full scope of the hyperactivity-impulsivity domain. Therefore, comparisons of these means cannot be reliably interpreted.

Our study's findings deviated from previous work identifying differences in ratings due to the type of measures being used and how the scores were obtained. Most of the studies comparing ADHD ratings between ethnic groups used other ADHD rating scales (e.g., ADHD-RS-IV, CRS, SNAP-IV) and sum/mean scores from these scales. However, there is limited research suggesting that these rating scales can be equivalently used across different ethnic groups and the use of sum/mean scores may be more susceptible to measurement biases (DuPaul et al., 2016; Miller et al., 2009; Stevanovic et al., 2017), so it is unclear whether the differences seen in ratings reflect *true* differences. Significant differences may exist when comparing sum/mean scores but not when comparing latent means due to potential biases in the measures[4]. Results from this current study help to support this given that parents of Asian and Hispanic/Latinx children may have different perceptions about hyperactive-impulsive behaviors despite having similar meanings attached to these constructs that may affect their scores. This highlights the need to further test for measurement invariance in other ADHD rating scales and to be mindful of different cultural interpretations of measures that might impact behavioral perceptions and endorsements before making conclusions about differences in ADHD symptoms or severity.

Clinical Implications

Results from this study provide important clinical implications regarding the use of the SDQ-HI for assessing and conceptualizing ADHD. The structure of the SDQ-HI appeared to align with current DSM conceptualizations of ADHD (i.e., inattentive and hyperactive/impulsive symptom domains; APA, 2022) and demonstrated configural and metric invariance across ethnicity. Such results suggested that clinicians may be able to deploy the measure as a screening tool for ADHD in real-world or research settings without needing to modify its content based on reporters' ethnicities. At the same time, caution may be needed when employing cut-off scores for the measure, particularly with respect to the hyperactive/impulsive symptom domain, as results supported the idea of ethnic variation in the way that symptoms within this domain are perceived by parents. Specifically, results suggested that although parents of children from different ethnic groups may attach similar meanings and interpretations of ADHD, clinicians

[4] Post-hoc analyses looking at differences in mean scores across ethnic groups were conducted (see Supplementary Material). Results indicated that White children had significantly higher inattention and hyperactivity-impulsivity mean scores than Hispanic/Latinx and Asian children.

may need to consider how culture might impact parents' perceptions of and tolerance toward behaviors that characterize the hyperactive/impulsive items, particularly in Hispanic/Latinx and Asian communities. In addition to asking clarifying questions to obtain more information on behavioral ratings, it may be important to ask questions about impairment, along with incorporating input from various other informants. Such methods may be important for fully understanding whether ADHD-related behaviors are clinically elevated or are a product of differing tolerance levels and/or response biases in parents with different ethnic backgrounds. Further, during feedback portions of assessments, it may be crucial for clinicians to spend extra time helping parents of Hispanic/Latinx and Asian children understand how ADHD is expressed using culturally relevant language.

Alongside this, findings may also inform efforts to increase ADHD-related service utilization in historically marginalized ethnic groups. Previous work has found that ethnically minoritized children are less likely to be identified and treated for ADHD (Davis et al., 2021; Yang et al., 2022; Wong & Landes, 2022). Several models have been created to help clarify pathways to service utilization, with one such model being the ADHD Help-Seeking Behavior Model (ADHD HSB Model; Eiraldi et al., 2002). This model suggests that children and their families may experience potential barriers to care across four stages of help-seeking behavior (i.e., Problem Recognition, Decision to Seek Help, Service Selection, Service Utilization; Eiraldi et al., 2002). The current study demonstrates that parents across ethnic groups interpret and rate inattentive behaviors similarly. As such, ethnically minoritized children are less likely to encounter barriers, with regards to parent ratings, in the Problem Recognition stage for inattentive behaviors. Additionally, findings from this study indicate that parents of Hispanic/Latinx and Asian children have lower tolerance for hyperactive behaviors and thus may be more likely to seek professional help because of this. However, as prior studies have shown that Black, Hispanic/Latinx, and Asian children are less likely to receive care (Davis et al., 2021; Yang et al., 2022), this indicates that these groups may still face cultural and societal barriers in other help-seeking stages. Other barriers include stigma, lower knowledge about ADHD, negative attitudes toward traditional mental health services, discriminatory practices, and cultural mismatch in care (Eiraldi et al., 2002). Thus, clinicians should not only be mindful of other barriers in other service utilization stages but also recognize that the types of barriers may differ based on cultural considerations. This awareness is important to potentially increase access to care and mitigate disparities in ADHD service use in ethnically minoritized populations.

Limitations and Future Research Directions

The study has many strengths, including the use of a nationally representative community sample of children, inclusive of diverse ethnic groups. However, despite these strengths and the importance of the current study in informing the assessment and treatment of ADHD and culturally sensitive procedures, several limitations need to be considered. First, as described earlier, latent means computed via Mplus do not account for non-invariant parameters. Additionally, as the SDQ-HI uses only five items, the removal of some items impacted our ability to effectively compare latent means, particularly for Hispanic/Latinx and Asian children. Therefore, future research should continue to examine for measurement invariance in other ADHD rating scales to determine whether true differences exist in these two ethnic groups' scores.

One other limitation of these results is that covariates, such as age and sex, were not included in the models. This could potentially affect some of these results given that prior research has shown sex and gender differences in ADHD symptoms and severity (Cherkasova et al., 2013; Loyer Carbonneau et al., 2021)[5]. The alignment optimization has recently allowed for the inclusion of covariates which may allow for further understanding of potential effects (Asparouhov & Muthén, 2023).

In addition to this, the current study was not able to include other variables to explain some of these results. More specifically, variables surrounding ADHD perceptions and cultural values represent a critical area of focus that future work may wish to expound upon.

Furthermore, other adult informants' reports (e.g., grandparent, aunty/uncle, sibling) were included, and informants' ethnic identities were unable to be parsed. Therefore, it is possible that the ethnicity of the informant may not necessarily match the child's ethnicity and may subsequently affect measurement interpretation and ratings due to their cultural values (Liang et al., 2016; Kang & Harvey, 2020). However, as the majority of the informants (93.5%) were the child's parents and a large proportion of children in the sample were of the same ethnicity as all parents in their household (84.6%), the ethnic match between the adult informant and child is relatively high (NCHS, 2020). Despite this, future research should attempt to account for the ethnic background of both the child and adult to interpret how results may be affected by the cultural context better.

Additionally, as the current study primarily includes parent reports in the home environment, there is limited understanding of the construct validity of the SDQ-HI for other informants, such as teachers. The teachers' ethnic background may interact with school environmental expectancies and lead to differential understandings and use of measures (Des Los Reyes & Kazdin, 2005; Pelham et al., 2005; Wexler et al., 2022). Thus, future research might

[5] CFA and alignment analyses were replicated using sex (male vs. female) and age (child vs. adolescent) as grouping variables (see Supplementary Material). Results suggested that the SDQ-HI was invariant across sex and age, with males and children (4-11) having higher inattention and hyperactivity-impulsivity than females and adolescents (12-17).

also extend this study by examining the construct validity of the SDQ-HI using teacher informant ratings across ethnicities. Given that ADHD symptoms need to be present in two or more settings to determine whether a child should be considered for a diagnosis (APA, 2022; Wolraich et al., 2019), further examining the construct validity of the SDQ-HI with teacher informants from various ethnic backgrounds may help to generalize its use and ensure a comprehensive understanding toward children's behaviors.

Lastly, due to limited sample sizes, children of other ethnic groups (e.g., American Indian/Alaska Native, Native Hawaiian/Pacific Islander), as well as multiethnic children were unable to be accounted for. Given that other ethnic groups may attach different meanings and perspectives (Gerdes et al., 2013; Liang et al., 2016), it is important to continually examine the construct validity of the SDQ-HI and do so before making comparisons between groups.

Conclusion

Given ethnic disparities in ADHD prevalence, along with negative outcomes associated with the disorder, measures that can appropriately screen for childhood ADHD across a range of ethnic groups are needed. The current study provides further support for the construct validity of the SDQ-HI, inclusive of measurement invariance in White, Black, Hispanic/Latinx, and Asian children. Therefore, the SDQ-HI may be a useful screener in the identification of ADHD, as it encapsulates two-dimensional conceptualizations of ADHD and is found to have equivalent structure and meaning across a range of ethnic groups. Additionally, this study provided an important step toward understanding differences in ADHD ratings. Despite prior studies indicating that there are ethnic differences in ADHD ratings, these results may be partially attributed to measurement biases. As found in this study, although ethnic groups had similar meanings and definitions of ADHD, differences may exist in endorsement and perception of behaviors. Clinicians should continue to consider how this might affect the assessment of ADHD, particularly regarding differing tolerance levels. Meanwhile, future research should continue to examine the measurement invariance of the SDQ-HI and other ADHD rating scales across the ethnic backgrounds of different informants and attempt to further understand ways in which culture plays a role in ADHD recognition. This study, along with other future studies, may help to provide new avenues to facilitate ADHD assessment and close the gap in the identification of ADHD across ethnic groups.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H.-C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*(3), 251–275. https://doi.org/10.1111/j.1469-7610.2007.01867.x

Achenbach T. M., Rescorla L.A. (2000). *Manual for the ASEBA school-age forms and profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.

Algorta, G. P., Dodd, A. L., Stringaris, A., & Youngstrom, E. A. (2016). Diagnostic efficiency of the SDQ for parents to identify ADHD in the UK: A ROC analysis. *European Child & Adolescent Psychiatry*, *25*(9), 949–957. https://doi.org/10.1007/s00787-015-0815-0

Alvarez, K., Fillbrunn, M., Green, J. G., Jackson, J. S., Kessler, R. C., McLaughlin, K. A., Sadikova, E., Sampson, N. A., & Alegría, M. (2019). Ethnicity, nativity, and lifetime risk of mental disorders in US adults. *Social Psychiatry and Psychiatric Epidemiology*, *54*(5), 553–565. https://doi.org/10.1007/s00127-018-1644-5

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised).

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.).

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

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.).

Arcia, E., & Fernandez, M. C. (1998). Cuban mothers' schemas of ADHD: Development, characteristics, and help seeking behavior. *Journal of Child and Family Studies*, *7*(3), 333–352. https://doi.org/10.1023/A:1022945628866

Arnett, A. B., Pennington, B. F., Willcutt, E. G., DeFries, J. C., & Olson, R. K. (2015). Sex differences in ADHD symptom severity. *Journal of Child Psychology and Psychiatry*, *56*(6), 632–639. https://doi.org/10.1111/jcpp.12337

Arnold, L. E., Elliott, M., Sachs, L., Bird, H., Kraemer, H. C., Wells, K. C., Abikoff, H. B., Comarda, A., Conners, C. K., Elliott, G. R., Greenhill, L. L., Hechtman, L., Hinshaw, S. P., Hoza, B., Jensen, P. S., March, J. S., Newcorn, J. H., Pelham, W. E., Severe, J. B., … Wigal, T. (2003). Effects of Ethnicity on Treatment Attendance, Stimulant Response/Dose, and 14-Month Outcome in ADHD. *Journal of Consulting and Clinical Psychology*, *71*(4), 713–727. https://doi.org/10.1037/0022-006X.71.4.713

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 495-508. https://doi.org/10.1080/10705511.2014.919210

Asparouhov, T., & Muthén, B. (2023). Multiple group alignment for exploratory and structural equation models. *Structural Equation Modeling*, *30*(2), 169–191. https://doi.org/10.1080/10705511.2022.2127100

Ault-Brutus, A., & Alegria, M. (2018). Ethnic differences in perceived need for mental health care and disparities in use of care among those with perceived need in 1990-1992 and 2001-2003. *Ethnicity & Health*, *23*(2), 142–157. https://doi.org/10.1080/13557858.2016.1252834

Arcia, E., & Fernández, M. C. (2003). From awareness to acknowledgment: The development of concern among Latina mothers of children with disruptive behaviors. *Journal of Attention Disorders*, *6*(4), 163-175. https://doi.org/10.1177/108705470300600403

Barbaresi, W. J., Colligan, R. C., Weaver, A. L., Voigt, R. G., Killian, J. M., & Katusic, S. K. (2013). Mortality, ADHD, and psychosocial adversity in adults with childhood ADHD: A Prospective Study. *Pediatrics* (Evanston), *131*(4), 637–644. https://doi.org/10.1542/peds.2012-2354

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

Bauermeister, J. J. (1992). Factor analyses of teacher ratings of attention-deficit hyperactivity and oppositional defiant symptoms in children aged four through thirteen years. *Journal of Clinical Child and Adolescent Psychology*, *21*(1), 27–34. https://doi.org/10.1207/s15374424jccp2101_5

Bauermeister, J. J., Alegría, M., Bird, H. R., Rubio-Stipec, M., & Canino, G. (1992). Are Attentional-Hyperactivity Deficits Unidimensional or Multidimensional Syndromes? Empirical Findings from a Community Survey. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*(3), 423–431. https://doi.org/10.1097/00004583-199205000-00007

Bauermeister, J. J., Canino, G., Bravo, M., Ramírez, R., Jensen, P. S., Chavez, L., Martíneztaboas, A., Ribera, J., Alegría, M., & García, P. (2003). Stimulant and psychosocial treatment of ADHD in Latino/Hispanic children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(7), 851–855. https://doi.org/10.1097/01.CHI.0000046864.56865.30

Bauermeister, J. J., Matos, M., Reina, G., Salas, C. C., Martínez, J. V., Cumba, E., & Barkley, R. A. (2005). Comparison of the DSM-IV combined and inattentive types of ADHD in a school-based sample of Latino/Hispanic children. *Journal of Child Psychology and Psychiatry*, *46*(2), 166–179. https://doi.org/10.1111/j.1469-7610.2004.00343.x

Biederman, J., Mick, E., & Faraone, S. V. (2000). Age-dependent decline of symptoms of attention deficit hyperactivity disorder: Impact of remission definition and symptom type. *The American Journal of Psychiatry*, *157*(5), 816–818. https://doi.org/10.1176/appi.ajp.157.5.816

Bitsko, R. H., Claussen, A. H., Lichstein, J., Black, L. I., Jones, S. E., Danielson, M. L., Hoenig, J. M., Davis Jack, S. P., Brody, D. J., Gyawali, S., Maenner, M. J., Warner, M., Holland, K. M., Perou, R., Crosby, A. E., Blumberg, S. J., Avenevoli, S., Kaminski, J. W., & Ghandour, R. M. (2022). Mental health surveillance among children - United States, 2013-2019. *Morbidity and Mortality Weekly Report*. *Supplement*, *71*(2), 1–42. https://doi.org/10.15585/mmwr.su7102a1

Bird, H. R., Canino, G. J., Davies, M., Ramírez, R., Chávez, L., Duarte, C., & Shen, S. (2005). The Brief Impairment Scale (BIS): A multidimensional scale of functional impairment for children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(7), 699–707. https://doi.org/10.1097/01.chi.0000163281.41383.94

Bird, H. R., Shaffer, D., Fisher, P., Gould, M. S., Staghezza, B., Chen, J. Y., & Hoven, C.

(1993). *Columbia Impairment Scale (CIS)* [Database record]. APA PsycTests. https://doi.org/10.1037/t06724-000

Bøe, T., Hysing, M., Skogen, J. C., & Breivik, K. (2016). The Strengths and Difficulties Questionnaire (SDQ): Factor structure and gender equivalence in Norwegian adolescents. *PloS One*, *11*(5), e0152202–e0152202. https://doi.org/10.1371/journal.pone.0152202

Bower, K. M., Thorpe, R. J., & LaVeist, T. A. (2013). Perceived racial discrimination and mental health in low-income, urban-dwelling whites. *International Journal of Health Services*, *43*(2), 267–280. https://doi.org/10.2190/HS.43.2.e

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., Visser, S., & Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics (Evanston), 127*(6), 1034–1042. https://doi.org/10.1542/peds.2010-2989

Bozinovic, K., McLamb, F., O'Connell, K., Olander, N., Feng, Z., Haagensen, S., & Bozinovic, G. (2021). U.S. national, regional, and state-specific socioeconomic factors correlate with child and adolescent ADHD diagnoses pre-COVID-19 pandemic. *Scientific Reports*, *11*(1), 22008–22008. https://doi.org/10.1038/s41598-021-01233-2

Braje, S. E., Murakami-Brundage, J., Hall, G. C. N., Wang, V. O., & Ge, X. (2012). Antisocial behavior and externalizing disorders among Asians, Asian Americans, and Pacific Islander Populations. In *Handbook of Adult Psychopathology in Asians*. Oxford University Press. https://doi.org/10.1093/med:psych/9780195179064.003.0010

Bruton, A., & Robles-Piña, R. A. (2009). Deficit thinking and Hispanic student achievement:

Scientific information resources. *Problems of Education in the 21st Century*, *15*, 41–48. http://dx.doi.org/10.3998/currents.17387731.0001.110

Bull, R., Lee, K., Koh, I. H. C., & Poon, K. K. L. (2016). Confirmatory factor analysis of the Strengths and Difficulties Questionnaire in Singaporean kindergartners. *Child: Care, Health & Development*, *42*(1), 109–116. https://doi.org/10.1111/cch.12288

Bussing, R., Fernandez, M., Harwood, M., Wei Hou, Garvan, C. W., Eyberg, S. M., & Swanson, J. M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: Psychometric properties and normative ratings from a school district sample. *Assessment (Odessa, Fla.)*, *15*(3), 317–328. https://doi.org/10.1177/1073191107313888

Bussing, R., Gary, F., Mills, T., & Garvan, C. (2003). Parental explanatory models of ADHD – Gender and cultural variations. *Social Psychiatry and Psychiatric Epidemiology*, *38*(10), 563–575. https://doi.org/10.1007/s00127-003-0674-8

Bussing, R., Mason, D. M., Bell, L., Porter, P., & Garvan, C. (2010). Adolescent Outcomes of Childhood Attention-Deficit/Hyperactivity Disorder in a Diverse Community Sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(6), 595–605. https://doi.org/10.1097/00004583-201006000-00008

Bussing, R., Schumann, E., Belin, T. R., Widawski, M., & Perwien, A. R. (1998). Diagnostic utility of two commonly used ADHD screening measures among special education students. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*(1), 74–82. https://doi.org/10.1097/00004583-199801000-00020

Canino, G., & Alegría, M. (2008). Psychiatric diagnosis–is it universal or relative to culture? *Journal of child psychology and psychiatry*, *49*(3), 237-250.

https://doi.org/10.1111/j.1469-7610.2007.01854.x

Cénat, Kokou-Kpolou, C. K., Blais-Rochette, C., Morse, C., Vandette, M.-P., Dalexis, R. D., Darius, W. P., Noorishad, P.-G., Labelle, P. R., & Kogan, C. S. (2022). Prevalence of ADHD among Black youth compared to White, Latino and Asian youth: A Meta-Analysis. *Journal of Clinical Child and Adolescent Psychology*, 1–16. https://doi.org/10.1080/15374416.2022.2051524

Centers for Disease Control and Prevention. (2022, August 9). Data and statistics about ADHD. Centers for Disease Control and Prevention. Retrieved June 26, 2023, from https://www.cdc.gov/ncbddd/adhd/data.html

Chang, L.-Y., Wang, M.-Y., & Tsai, P.-S. (2016). Diagnostic accuracy of rating scales for attention-deficit/hyperactivity disorder: A meta-analysis. *Pediatrics (Evanston), 137*(3), e20152749–e20152749. https://doi.org/10.1542/peds.2015-2749

Chen, F. F., & West, S. G. (2008). Measuring individualism and collectivism: The importance of considering differential components, reference groups, and measurement invariance. *Journal of Research in Personality*, *42*(2), 259–294. https://doi.org/10.1016/j.jrp.2007.05.006

Cherkasova, M., Sulla, E. M., Dalena, K. L., Pondé, M. P., & Hechtman, L. (2013). Developmental course of attention deficit hyperactivity disorder and its predictors. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *22*(1), 47–54.

Chhabildas, N., Pennington, B. F., & Willcutt, E. G. (2001). A comparison of the neuropsychological profiles of the DSM-IV subtypes of ADHD. *Journal of Abnormal Child Psychology*, *29*(6), 529–540. https://doi.org/10.1023/A:1012281226028

Chiorri, C., Hall, J., Casely-Hayford, J., & Malmberg, L.-E. (2016). Evaluating measurement invariance between parents using the Strengths and Difficulties Questionnaire (SDQ). *Assessment (Odessa, Fla.)*, *23*(1), 63–74. https://doi.org/10.1177/1073191114568301

Chung, Jiang, S.-F., Paksarian, D., Nikolaidis, A., Castellanos, F. X., Merikangas, K. R., & Milham, M. P. (2019). Trends in the prevalence and incidence of attention-deficit/hyperactivity disorder among adults and children of different ethnic groups. *JAMA Network Open*, *2*(11), e1914344–e1914344. https://doi.org/10.1001/jamanetworkopen.2019.14344

Coker, Elliott, M. N., Toomey, S. L., Schwebel, D. C., Cuccaro, P., Emery, S. T., Davies, S. L., Visser, S. N., & Schuster, M. A. (2016). Ethnic disparities in ADHD diagnosis and treatment. *Pediatrics (Evanston), 138*(3), e20160407–e20160407. https://doi.org/10.1542/peds.2016-0407

Collett, B. R., Ohan, J. L., & Myers, K. M. (2003). Ten-year review of rating scales. V: Scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(9), 1015–1037. https://doi.org/10.1097/01.CHI.0000070245.24125.B6

Conners, C. (1997). *Conners' Rating Scales-Revised Technical* Manual. North Tonawanda, NY: Multi-Health Systems.

Cook, B. L., Trinh, N.-H., Li, Z., Hou, S. S.-Y., & Progovac, A. M. (2017). Trends in racial-ethnic disparities in access to mental health care, 2004–2012. Psychiatric Services (Washington, D.C.), 68(1), 9–16. https://doi.org/10.1176/appi.ps.201500453

Croft, S., Stride, C., Maughan, B., & Rowe, R. (2015). Validity of the strengths and difficulties questionnaire in preschool-aged children. *Pediatrics (Evanston)*, *135*(5), e1210–e1219. https://doi.org/10.1542/peds.2014-2920

Cuffe, Moore, C. G., & McKeown, R. E. (2005). Prevalence and correlates of ADHD symptoms in the national health interview survey. *Journal of Attention Disorders*, *9*(2), 392–401. https://doi.org/10.1177/1087054705280413

Curatolo, P., D'Agati, E., & Moavero, R. (2010). The neurobiological basis of ADHD. *Italian Journal of Pediatrics*, *36*(1), 79–79. https://doi.org/10.1186/1824-7288-36-79

Davis, J. K., Feygin, Y., Creel, L., Kong, M., Sun, J., Lohr, W. D., Williams, P. G., Le, J., Jones, V. F., Trace, M., & Pasquenza, N. (2021). Disparities in ADHD diagnosis and treatment by ethnicity in youth receiving Kentucky medicaid in 2017. *Ethnicity & Disease*, *31*(1), 67–76. https://doi.org/10.18865/ed.31.1.67

Davis, R. E., Resnicow, K., & Couper, M. P. (2011). Survey response styles, acculturation, and culture among a sample of Mexican American adults. *Journal of Cross-Cultural Psychology*, *42*(7), 1219–1236. https://doi.org/10.1177/0022022110383317

Davison, J. C., & Ford, D. Y. (2001). Perceptions of attention deficit hyperactivity disorder in one African American community. *The Journal of Negro Education*, *70*(4), 264–274. https://doi.org/10.2307/3211279

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858–900. https://doi.org/10.1037/a0038498

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), 483–509. https://doi.org/10.1037/0033-2909.131.4.483

Dickey, W. C., & Blumberg, S. J. (2004). Revisiting the factor structure of the Strengths and Difficulties Questionnaire: United States, 2001. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*(9), 1159–1167. https://doi.org/10.1097/01.chi.0000132808.36708.a9

Dietrich, S., & Hernandez, E. (2022, December 13). *Nearly 68 million people spoke a language other than English at home in 2019*. U.S. Census Bureau. https://www.census.gov/library/stories/2022/12/languages-we-speak-in-united-states.html

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Reid, R., McGoey, K. E., & Ikeda, M. J. (1997). Teacher ratings of attention deficit hyperactivity disorder symptoms: Factor structure and normative data. *Psychological Assessment*, *9*, 436–444. http://dx.doi.org/10.1037/1040-3590.9.4.436

DuPaul, Anastopoulos, A., Power, T., Reid, R., Ikeda, M., & McGoey, K. (1998). Parent ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Journal of Psychopathology and Behavioral Assessment*, *20*(1), 83–102. https://doi.org/10.1023/A:1023087410712

DuPaul G. J., Power T. J., Anastopoulos A. D., Reid R. (1998). *ADHD Rating Scale–IV: Checklists, norms, and clinical interpretation*. New York, NY: Guilford Press.

DuPaul, Reid, R., Anastopoulos, A. D., Lambert, M. C., Watkins, M. W., & Power, T. J. (2016).

Parent and teacher ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Psychological Assessment*, *28*(2), 214–225. https://doi.org/10.1037/pas0000166

Elander, J., & Rutter, M. (1996). Use and development of the rutter parents' and teachers' and teachers' scales. *International Journal of Methods in Psychiatric Research*, *6*(2), 63–78. https://doi.org/10.1002/(SICI)1234-988X(199607)6:2<63::AID-MPR151>3.3.CO;2-M

Epstein, J. N., & Loren, R. E. (2013). Changes in the definition of ADHD in DSM-5: subtle but important. *Neuropsychiatry*, *3*(5), 455–458. https://doi.org/10.2217/npy.13.59

Epstein, J. N., Willoughby, M., Valencia, E. Y., Tonev, S. T., Abikoff, H. B., Arnold, L. E., & Hinshaw, S. P. (2005). The role of children's ethnicity in the relationship between teacher ratings of attention-deficit/hyperactivity disorder and observed classroom behavior. *Journal of Consulting and Clinical Psychology*, *73*(3), 424–434. https://doi.org/10.1037/0022-006X.73.3.424

Evans, S. W., Brady, C. E., Harrison, J. R., Bunford, N., Kern, L., State, T., & Andrews, C. (2013). Measuring ADHD and ODD symptoms and impairment using high school teachers' ratings. *Journal of Clinical Child and Adolescent Psychology*, *42*(2), 197–207. https://doi.org/10.1080/15374416.2012.738456

Fairman, K. A., Peckham, A. M., & Sclar, D. A. (2020). Diagnosis and treatment of ADHD in the United States: Update by gender and race. *Journal of Attention Disorders*, *24*(1), 10–19. https://doi.org/10.1177/1087054716688534

Fiscella, K., Franks, P., Gold, M. R., & Clancy, C. M. (2000). Inequality in quality: Addressing socioeconomic, racial, and ethnic disparities in health care. *JAMA: The Journal of the American Medical Association*, *283*(19), 2579–2584. https://doi.org/10.1001/jama.283.19.2579

Flanagin, A., Frey, T., & Christiansen, S. L. (2021). Updated guidance on the reporting of race and ethnicity in medical and science journals. *JAMA: The Journal of the American Medical Association*, *326*(7), 621–627. https://doi.org/10.1001/jama.2021.13304

Fleming, M., Fitton, C. A., Steiner, M. F. C., McLay, J. S., Clark, D., King, A., Mackay, D. F., & Pell, J. P. (2017). Educational and health outcomes of children treated for attention-deficit/hyperactivity disorder. *JAMA Pediatrics*, *171*(7), e170691–e170691. https://doi.org/10.1001/jamapediatrics.2017.0691

Flory, K., Shi, D., Siceloff, E. R., Roberts, A. M., Castellanos, R., Neger, E., Taylor, S., & Benson, K. (2021). The factor structure and gender invariance of ADHD symptoms in college students. *Assessment (Odessa, Fla.)*, *28*(1), 57–72. https://doi.org/10.1177/1073191120918934

Gadow, K. D., & Sprafkin, J. (1994). *Child symptom inventories manual*. Stony Brook, NY: Checkmate Plus.

Galvez-Contreras, A. Y., Vargas-de la Cruz, I., Beltran-Navarro, B., Gonzalez-Castaneda, R. E., & Gonzalez-Perez, O. (2022). Therapeutic approaches for ADHD by developmental stage and clinical presentation. *International Journal of Environmental Research and Public Health*, *19*(19), 12880-12896. https://doi.org/10.3390/ijerph191912880

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

Gerdes, A. C., Lawton, K. E., Haack, L. M., & Hurtado, G. D. (2013). Assessing ADHD in Latino families: Evidence for moving beyond symptomatology. *Journal of Attention Disorders*, *17*(2), 128–140. https://doi.org/10.1177/1087054711427396

Getahun, Jacobsen, S. J., Fassett, M. J., Chen, W., Demissie, K., & Rhoads, G. G. (2013). Recent trends in childhood attention-deficit/hyperactivity disorder. *JAMA Pediatrics*, *167*(3), 1–7. https://doi.org/10.1001/2013.jamapediatrics.401

Gomez, R., & Stavropoulos, V. (2021). Confirmatory factor analysis and exploratory structural equation modeling of the structure of attention-deficit/hyperactivity disorder symptoms in adults. *Assessment*, *28*(6), 1570-1582. https://doi.org/10.1177/1073191120905892

Gómez-Benito, J., Van de Vijver, F. J., Balluerka, N., & Caterino, L. (2019). Cross-cultural and gender differences in ADHD among young adults. *Journal of attention disorders*, *23*(1), 22-31. https://doi.org/10.1177/1087054715611748

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(11), 1337–1345. https://doi.org/10.1097/00004583-200111000-00015

Goodman, R., Meltzer, H., & Bailey, V. (1998). The strengths and difficulties questionnaire: A pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry*, *7*(3), 125–130. https://doi.org/10.1007/s007870050057

Goodman, R., & Scott, S. (1999). Comparing the strengths and difficulties questionnaire and the child behavior checklist : Is small beautiful? *Journal of Abnormal Child Psychology*, *27*(1), 17–24. https://doi.org/10.1023/A:1022658222914

Graybill, E., Roach, A., & Barger, B. (2021). Factor structure of the self-report strength and difficulties questionnaire in a diverse U.S. sample. *Journal of Psychopathology and Behavioral Assessment*, *43*(2), 388–398. https://doi.org/10.1007/s10862-020-09833-4

Haack, L. M., Meza, J., Jiang, Y., Araujo, E. J., & Pfiffner, L. (2018). Influences to ADHD problem recognition: Mixed-method investigation and recommendations to reduce disparities for Latino youth. *Administration and Policy in Mental Health and Mental Health Services Research*, *45*(6), 958–977. https://doi.org/10.1007/s10488-018-0877-7

Hall, C. L., Guo, B., Valentine, A. Z., Groom, M. J., Daley, D., Sayal, K., & Hollis, C. (2019). The validity of the Strengths and Difficulties Questionnaire (SDQ) for children with ADHD symptoms. *PloS One*, *14*(6), e0218518–e0218518. https://doi.org/10.1371/journal.pone.0218518

Harvey, E. A., Fischer, C., Weieneth, J. L., Hurwitz, S. D., & Sayer, A. G. (2013). Predictors of discrepancies between informants' ratings of preschool-aged children's behavior: An examination of ethnicity, child characteristics, and family functioning. *Early Childhood Research Quarterly*, *28*(4), 668–682. https://doi.org/10.1016/j.ecresq.2013.05.002

Hassan, N. M., Husain, R., Aziz, A. A., Mazubir, N. N., Juhari, S. N., & Daud, N. (2019). Validation of the Malay Version of the self-rated strength and difficulties (SDQ) questionnaire. *International Journal of Academic Research in Business and Social Sciences*, *9*(11). https://doi.org/10.6007/IJARBSS/v9-i11/6553

Hawes, D. J., & Dadds, M. R. (2004). Australian data and psychometric properties of the

Strengths and Difficulties Questionnaire. *Australian and New Zealand Journal of Psychiatry*, *38*(8), 644–651. https://doi.org/10.1080/j.1440-1614.2004.01427.x

Hays, D. G., Prosek, E. A., & McLeod, A. L. (2010). A mixed methodological analysis of the role of culture in the clinical decision-making process. *Journal of Counseling and Development*, *88*(1), 114–121. https://doi.org/10.1002/j.1556-6678.2010.tb00158.x

He, J. P., Burstein, M., Schmitz, A., & Merikangas, K. R. (2013). The Strengths and Difficulties Questionnaire (SDQ): The factor structure and scale validation in U.S. adolescents. *Journal of Abnormal Child Psychology*, *41*(4), 583–595. https://doi.org/10.1007/s10802-012-9696-6

Hillemeier, M. M., Foster, E. M., Heinrichs, B., & Heier, B. (2007). Racial differences in parental reports of attention-deficit/hyperactivity disorder Behaviors. *Journal of Developmental and Behavioral Pediatrics*, *28*(5), 353–361. https://doi.org/10.1097/DBP.0b013e31811ff8b8

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jang, S. J. (2002). Race, ethnicity, and deviance: A study of Asian and non-Asian adolescents in America. *Sociological Forum (Randolph, N.J.)*, *17*(4), 647–680. https://doi.org/10.1023/A:1021081524775

Jimenez, D. E., Alegría, M., Chen, C., Chan, D., & Laderman, M. (2010). Prevalence of psychiatric illnesses in older ethnic minority adults. *Journal of the American Geriatrics Society (JAGS)*, *58*(2), 256–264. https://doi.org/10.1111/j.1532-5415.2009.02685.x

Jose, P. E., Huntsinger, C. S., Huntsinger, P. R., & Liaw, F.-R. (2000). Parental values and practices relevant to young children's social development in Taiwan and the United States. *Journal of Cross-Cultural Psychology*, *31*(6), 677–702. https://doi.org/10.1177/0022022100031006002

Kang, S., & Harvey, E. A. (2020). Racial differences between Black parents' and White teachers' perceptions of attention-deficit/hyperactivity disorder behavior. *Journal of Abnormal Child Psychology*, *48*(5), 661–672. https://doi.org/10.1007/s10802-019-00600-y

Karlsson, P., Larm, P., Svensson, J., & Raninen, J. (2022). The factor structure of the Strength and Difficulties Questionnaire in a national sample of Swedish adolescents: Comparing 3 and 5-factor models. *Plos one*, *17*(3), e0265481.

Kiely, B., & Adesman, A. (2015). What we do not know about ADHD… yet. *Current Opinion in Pediatrics*, *27*(3), 395–404. https://doi.org/10.1097/MOP.0000000000000229

Kjærandsen, K. S., Handegård, B. H., Brøndbo, P. H., & Halvorsen, M. B. (2023). Impact supplement of the strengths and difficulties questionnaire in the assessment of functional impairment in children with ADHD or ASD in a mixed neuropediatric sample: A partial validation study. *Journal of Mental Health Research in Intellectual Disabilities*, *16*(1), 1–22. https://doi.org/10.1080/19315864.2022.2148788

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Koutsoklenis, A., & Honkasilta, J. (2022). ADHD in the DSM-5-TR: What has changed and what has not. *Frontiers in Psychiatry*, *13*, 1064141–1064141. https://doi.org/10.3389/fpsyt.2022.1064141

Kovacs, S., & Sharp, C. (2014). Criterion validity of the Strengths and Difficulties Questionnaire

(SDQ) with inpatient adolescents. *Psychiatry Research*, *219*(3), 651–657. https://doi.org/10.1016/j.psychres.2014.06.019

Kuntsi, J., Pinto, R., Price, T. S., van der Meere, J., Frazier-Wood, A. C., & Asherson, P. (2014). The separation of ADHD inattention and hyperactivity-impulsivity symptoms: pathways from genetic effects to cognitive impairments and symptoms. *Journal of Abnormal Child Psychology*, *42*(1), 127–136. https://doi.org/10.1007/s10802-013-9771-7

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

Lahey, B. B., Pelham, W. E., Schaughency, E. A., Atkins, M. S., Murphy, H. A., Hynd, G., Russo, M., Hartdagen, S., & Lorys-Vernon, A. (1988). Dimensions and Types of Attention Deficit Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*(3), 330–335. https://doi.org/10.1097/00004583-198805000-00011

Langberg, J. M., Froehlich, T. E., Loren, R. E., Martin, J. E., & Epstein, J. N. (2008). Assessing children with ADHD in primary care settings. *Expert Review of Neurotherapeutics*, *8*(4), 627–641. https://doi.org/10.1586/14737175.8.4.627

Langsdorf, R., Anderson, R. P., Waechter, D., Madrigal, J. F., & Juarez, L. J. (1979). Ethnicity, social class, and perception of hyperactivity. *Psychology in the Schools*, *16*(2), 293–298. https://doi.org/10.1002/1520-6807(197904)16:2<293::AID-PITS2310160221>3.0.CO;2-1

Larsson, H., Lichtenstein, P., & Larsson, J. O. (2006). Genetic Contributions to the Development of ADHD Subtypes from Childhood to Adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(8), 973–981. https://doi.org/10.1097/01.chi.0000222787.57100.d8

Lau, M., Lin, H., & Flores, G. (2012). Ethnic disparities in health and health care among U.S. adolescents. *Health Services Research*, *47*(5), 2031–2059. https://doi.org/10.1111/j.1475-6773.2012.01394.x

Lawson, G. M., Nissley-Tsiopinis, J., Nahmias, A., McConaughy, S. H., & Eiraldi, R. (2017). Do parent and teacher-report of ADHD symptoms in children differ by SES and racial status? *Journal of Psychopathology and Behavioral Assessment*, *39*(3), 426–440. https://doi.org/10.1007/s10862-017-9591-0

Lee, S. J. (1994). Behind the model-minority stereotype: Voices of high- and low-achieving Asian American students. *Anthropology & Education Quarterly*, *25*(4), 413–429. https://doi.org/10.1525/aeq.1994.25.4.04x0530j

Liang, J., Matheson, B. E., & Douglas, J. M. (2016). Mental health diagnostic considerations in ethnic minority youth. *Journal of Child and Family Studies*, *25*(6), 1926–1940. https://doi.org/10.1007/s10826-015-0351-z

Liu, M., Conrad, F. G., & Lee, S. (2017). Comparing acquiescent and extreme response styles in face-to-face and web surveys. *Quality & Quantity*, *51*(2), 941–958. https://doi.org/10.1007/s11135-016-0320-7

London, A. S., & Landes, S. D. (2016). Attention deficit hyperactivity disorder and adult mortality. *Preventive Medicine*, *90*, 8–10. https://doi.org/10.1016/j.ypmed.2016.06.021

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

Lu, C., Ahmed, R., Lamri, A., & Anand, S. S. (2022). Use of race, ethnicity, and ancestry data in health research. *PLOS Global Public Health*, *2*(9). https://doi.org/10.1371/journal.pgph.0001060

Lynch, A., & Davison, K. (2022). Gendered expectations on the recognition of ADHD in young women and educational implications. *Irish Educational Studies*, *ahead-of-print*(ahead-of-print), 1–21. https://doi.org/10.1080/03323315.2022.2032264

Mak, W., & Rosenblatt, A. (2002). Demographic influences on psychiatric diagnoses among youth served in California systems of care. *Journal of Child and Family Studies*, *11*(2), 165–178. https://doi.org/10.1023/A:1015173508474

Martel, M. M. (2013). Sexual selection and sex differences in the prevalence of childhood externalizing and adolescent internalizing disorders. *Psychological Bulletin*, *139*(6), 1221–1259. https://doi.org/10.1037/a0032247

Marti, F. A., Pourat, N., Lee, C., & Zima, B. T. (2022). Overview of ten child mental health clinical outcome measures: Testing of psychometric properties with diverse client populations in the U.S. *Administration and Policy in Mental Health and Mental Health Services Research*, *49*(2), 197–225. https://doi.org/10.1007/s10488-021-01157-z

Miller, T. W., Nigg, J. T., & Miller, R. L. (2009). Attention deficit hyperactivity disorder in African American children: What can be concluded from the past ten years? *Clinical Psychology Review*, *29*(1), 77–86. https://doi.org/10.1016/j.cpr.2008.10.001

Morgan, Li, H., Cook, M., Farkas, G., Hillemeier, M. M., & Lin, Y. (2016). Which kindergarten children are at greatest risk for attention-deficit/hyperactivity and conduct disorder symptomatology as adolescents? *School Psychology Quarterly*, *31*(1), 58–75. https://doi.org/10.1037/spq0000123

Mulraney, M., Arrondo, G., Musullulu, H., Iturmendi-Sabater, I., Cortese, S., Westwood, S. J., Donno, F., Banaschewski, T., Simonoff, E., Zuddas, A., Döpfner, M., Hinshaw, S. P., & 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 and Adolescent Psychiatry*, *61*(8), 982–996. https://doi.org/10.1016/j.jaac.2021.11.031

Mulvaney-Day, N., Marshall, T., Downey Piscopo, K., Korsen, N., Lynch, S., Karnell, L. H., Moran, G. E., Daniels, A. S., & Ghose, S. S. (2018). Screening for behavioral health conditions in primary care settings: A systematic review of the literature. *Journal of General Internal Medicine: JGIM*, *33*(3), 335–346. https://doi.org/10.1007/s11606-017-4181-0

Muthén, L.K. and Muthén, B.O. (1998-2023). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

National Center for Health Statistics (2020). *Survey description, National Health Interview Survey, 2019*. U.S. Department of Health and Human Services Centers for Disease Control and Prevention. https://nhis.ipums.org/nhis/resources/srvydesc2019.pdf

Neal, L. V. I., McCray, A. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The effects of African American movement styles on teachers' perceptions and reactions. *The Journal of Special Education*, *37*(1), 49–57. https://doi.org/10.1177/00224669030370010501

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European*

Journal of Social Psychology, *15*(3), 263–280. https://doi.org/10.1002/ejsp.2420150303

Nielsen, L. G., Rimvall, M. K., Clemmensen, L., Munkholm, A., Elberling, H., Olsen, E. M., Rask, C. U., Skovgaard, A. M., & Jeppesen, P. (2019). The predictive validity of the Strengths and Difficulties Questionnaire in preschool age to identify mental disorders in preadolescence. *PloS One*, *14*(6), e0217707–e0217707. https://doi.org/10.1371/journal.pone.0217707

Pastor, P. N., & Reuben, C. A. (2005). Ethnic differences in ADHD and LD in young school-age children: Parental reports in the national health interview survey. *Public Health Reports (1974)*, *120*(4), 383–392. https://doi.org/10.1177/003335490512000405

Patil, P. A., Porche, M. V., Shippen, N. A., Dallenbach, N. T., & Fortuna, L. R. (2018). Which girls, which boys? The intersectional risk for depression by race and ethnicity, and gender in the US. *Clinical psychology review*, *66*, 51-68. https://doi.org/10.1016/j.cpr.2017.12.003

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

Pham, A. V. (2013). Attention deficit hyperactivity disorder among Asian American families: Challenges in assessment and treatment. *Communiqué (National Association of School Psychologists)*, *41*(6), 9–10.

Phinney, J. S. (1996). When we talk about American ethnic groups, what do we mean? *The American Psychologist*, *51*(9), 918–927. https://doi.org/10.1037/0003-066X.51.9.918

Pillow, D. R., Pelham, W. E., Hoza, B., Molina, B. S. G., & Stultz, C. H. (1998). Confirmatory factor analyses examining attention deficit hyperactivity disorder symptoms and other childhood disruptive behaviors. *Journal of Abnormal Child Psychology*, *26*(4), 293–309. https://doi.org/10.1023/A:1022658618368

Posserud, M.-B., Ullebø, A. K., Plessen, K. J., Stormark, K. M., Gillberg, C., & Lundervold, A. J. (2014). Influence of assessment instrument on ADHD diagnosis. *European Child & Adolescent Psychiatry*, *23*(4), 197–205. https://doi.org/10.1007/s00787-013-0442-6

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Nguyen, L., Huang, L. N., Arganza, G. F., & Liao, Q. (2007). The influence of ethnicity on psychiatric diagnoses and clinical characteristics of children and adolescents in children's services. *Cultural Diversity & Ethnic Minority Psychology*, *13*(1), 18–25. https://doi.org/10.1037/1099-9809.13.1.18

Office of Management and Budget. (1997). Revisions to the standards for the classification of federal data on ethnicity. *Federal Register*, *62*(210), 58782-58790.

Øvergaard, K. R., Madsen, K. B., Oerbeck, B., Friis, S., & Obel, C. (2019). The predictive validity of the Strengths and Difficulties Questionnaire for child attention-deficit/hyperactivity disorder. *European Child & Adolescent Psychiatry*, *28*(5), 625–633. https://doi.org/10.1007/s00787-018-1226-9

Øvergaard, K. R., Oerbeck, B., Friis, S., Pripp, A. H., Biele, G., Aase, H., & Zeiner, P. (2018). Attention-deficit/hyperactivity disorder in preschoolers: The accuracy of a short screener. *Journal of the American Academy of Child and Adolescent Psychiatry*, *57*(6), 428–435. https://doi.org/10.1016/j.jaac.2018.03.008

Owens, J. (2021). Parental intervention in school, academic pressure, and childhood diagnoses of ADHD. *Social Science & Medicine (1982)*, *272*, 113746–113746. https://doi.org/10.1016/j.socscimed.2021.113746

Quebles, Solomon, O., Smith, K. A., Rao, S. R., Lu, F., Azen, C., Anaya, G., & Yin, L. (2020). Ethnic differences in behavioral problems and medication use among children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, *125*(5), 369–388. https://doi.org/10.1352/1944-7558-125.5.369

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

Ramirez, R., & Shapiro, E. (1998). Teacher ratings of attention deficit hyperactivity disorder symptoms in Hispanic children. *Journal of Psychopathology and Behavioral Assessment*, *20*(4), 275–293. https://doi.org/10.1023/A:1021934705392

Reid, R. (1995). Assessment of ADHD with culturally different groups: The use of behavioral rating scales. *School Psychology Review*, *24*(4), 537–560. https://doi.org/10.1080/02796015.1995.12085787

Reid, R., Casat, C., Norton, H., Anastopoulos, A., & Temple, E. (2001). Using behavior rating scales for ADHD across ethnic groups: The IOWA conners. *Journal of Emotional and Behavioral Disorders*, *9*(4), 210–218. https://doi.org/10.1177/106342660100900401

Reid, R., DuPaul, G., Power, T., Anastopoulos, A., Rogers-Adkinson, D., Noll, M., & Riccio, C. (1998). Assessing culturally different students for attention deficit hyperactivity disorder using behavior rating scales. *Journal of Abnormal Child Psychology*, *26*(3), 187–198. https://doi.org/10.1023/A:1022620217886

Reynolds, C. R., & Kamphaus R. W. (1992). *Behavior assessment system for children*. Circle Pines, MN: AGS.

Rhee, S. H., Waldman, I. D., Hay, D. A., & Levy, F. (1999). Sex differences in genetic and environmental influences on DSM–III–R attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology*, *108*(1), 24–41. https://doi.org/10.1037/0021-843X.108.1.24

Rimvall, M. K., Elberling, H., Rask, C. U., Helenius, D., Skovgaard, A. M., & Jeppesen, P. (2014). Predicting ADHD in school age when using the Strengths and Difficulties Questionnaire in preschool age: A longitudinal general population study, CCC2000. *European Child & Adolescent Psychiatry*, *23*(11), 1051–1060. https://doi.org/10.1007/s00787-014-0546-7

Rogge, J., Koglin, U., & Petermann, F. (2018). Do they rate in the same way?: Testing of measurement invariance across parent and teacher SDQ ratings. *European Journal of Psychological Assessment: Official Organ of the European Association of Psychological Assessment*, *34*(2), 69–78. https://doi.org/10.1027/1015-5759/a000445

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

Scheres, A., Lee, A., & Sumiya, M. (2008). Temporal reward discounting and ADHD: Task and symptom specific effects. *Journal of Neural Transmission*, *115*(2), 221–226. https://doi.org/10.1007/s00702-007-0813-6

Shi, Hunter Guevara, L. R., Dykhoff, H. J., Sangaralingham, L. R., Phelan, S., Zaccariello, M. J., & Warner, D. O. (2021). Racial dispairites in diagnosis of attention-deficity/hyperactivity disorder in a US national birth cohort. *JAMA Network Open*, *4*(3), e210321–e210321.https://doi.org/10.1001/jamanetworkopen.2021.0321

Sonuga-Barke, E. J. S. (2002). Psychological heterogeneity in AD/HD: A dual pathway model of behaviour and cognition: Neurobehavioural mechanisms in ADHD. *Behavioural Brain Research*, *130*(1-2), 29–36. https://doi.org/10.1016/s0166-4328(01)00432-6

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

Spence, N. D., Wells, S., Graham, K., & George, J. (2016). Racial discrimination, cultural resilience, and stress. *Canadian Journal of Psychiatry*, *61*(5), 298–307. https://doi.org/10.1177/0706743716638653

Spring, Blunden, D., Greenberg, L. M., & Yellin, A. M. (1977). Validity and norms of a hyperactivity rating scale. *The Journal of Special Education*, *11*(3), 313–321. https://doi.org/10.1177/002246697701100306

Staatz, B. C., Kelly, Y., Lacey, R. E., & Hardy, R. (2023). Investigating the factorial structure and measurement invariance of the parent-reported strengths and difficulties questionnaire at 11 years of age from the UK Millennium Cohort Study. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-023-02156-1

Stevanovic, D., Jafari, P., Knez, R., Franic, T., Atilola, O., Davidovic, N., Bagheri, Z., & Lakic, A. (2017). Can we really use available scales for child and adolescent psychopathology across cultures? A systematic review of cross-cultural measurement invariance data. *Transcultural Psychiatry*, *54*(1), 125–152. https://doi.org/10.1177/1363461516689215

Stolk, Y., Kaplan, I., & Szwarc, J. (2017). Review of the strengths and difficulties questionnaire translated into languages spoken by children and adolescents of refugee background. *International Journal of Methods in Psychiatric Research*, *26*(4), e1568–n/a. https://doi.org/10.1002/mpr.1568

Stone, L. L., Otten, R., Engels, R. C. M. E., Vermulst, A. A., & Janssens, J. M. A. M. (2010). Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year-olds: A review. *Clinical Child and Family Psychology Review*, *13*(3), 254–274. https://doi.org/10.1007/s10567-010-0071-2

Swanson, J. (1992). *School based assessments and interventions for ADD students*. Irvine, CA: KC.

Takeda, T., Nissley-Tsiopinis, J., Nanda, S., & Eiraldi, R. (2020). Factors associated with discrepancy in parent–teacher-reporting of symptoms of ADHD in a large clinic-referred sample of children. *Journal of Attention Disorders*, *24*(11), 1605–1615. https://doi.org/10.1177/1087054716652476

Taylor, E., Schachar, R., Thorley, G., & Wieselberg, M. (1986). Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients. *British Journal of Psychiatry*, *149*(6), 760–767. https://doi.org/10.1192/bjp.149.6.760

Vilsaint, C. L., NeMoyer, A., Fillbrunn, M., Sadikova, E., Kessler, R. C., Sampson, N. A., Alvarez, K., Green, J. G., McLaughlin, K. A., Chen, R., Williams, D. R., Jackson, J. S., & Alegría, M. (2019). Ethnic differences in 12-month prevalence and persistence of mood, anxiety, and substance use disorders: Variation by nativity and socioeconomic status. *Comprehensive Psychiatry*, *89*, 52–60. https://doi.org/10.1016/j.comppsych.2018.12.008

Ullebø, A. K., Posserud, M.-B., Heiervang, E., Gillberg, C., & Obel, C. (2011). Screening for the attention deficit hyperactivity disorder phenotype using the strength and difficulties questionnaire. *European Child & Adolescent Psychiatry*, *20*(9), 451–458. https://doi.org/10.1007/s00787-011-0198-9

Villodas, Pfiffner, L. J., Moses, J. O., Hartung, C., & McBurnett, K. (2019). The roles of student gender, race, and psychopathology in teachers' identification of students for services. *Children and Youth Services Review*, *107*, 104468–. https://doi.org/10.1016/j.childyouth.2019.104468

Wang, R., Hempton, B., Dugan, J. P., & Komives, S. R. (2008). Cultural differences: Why do Asians avoid extreme responses? *Survey Practice*, *1*(3), 1–7. https://doi.org/10.29115/SP-2008-0011

Watabe, Y., Lee, S. K., & Matsuhashi, Y. (2024). Cross-cultural investigation in differential perceptions of externalizing symptoms. *Japanese Psychological Research*, *66*(3), 255–263. https://doi.org/10.1111/jpr.12411

Wells, E. A., Morrison, D. M., Gillmore, M. R., Catalano, R. F., Iritani, B., & Hawkins, J. D. (1992). Race differences in antisocial behaviors and attitudes and early initiation of substance use. *Journal of Drug Education*, *22*(2), 115–130. https://doi.org/10.2190/3BHH-3NAT-BYNK-D3VC

Wexler, D., Salgado, R., Gornik, A., Peterson, R., & Pritchard, A. (2022). What's race got to do with it? Informant rating discrepancies in neuropsychological evaluations for children with ADHD. *Clinical Neuropsychologist*, *36*(2), 264–286. https://doi.org/10.1080/13854046.2021.1944671

Wolraich, M., Feurer, I., Hannah, J., Baumgaertel, A., & Pinnock, T. (1998). Obtaining systematic teacher-reports of disruptive behavior disorders utilizing DSM-IV. *Journal of Abnormal Child Psychology*, *26*(2), 141–152. https://doi.org/10.1023/A:1022673906401

Woerner, W., Fleitlich-Bilyk, B., Martinussen, R., Fletcher, J., Cucchiaro, G., Dalgalarrondo, P., Lui, M., & Tannock, R. (2004). The strengths and difficulties questionnaire overseas: evaluations and applications of the SDQ beyond Europe. *European Child & Adolescent Psychiatry*, *13 Suppl 2*(S2), II47–54. https://doi.org/10.1007/s00787-004-2008-0

Wolraich, M. L., Hagan, J., Allan, C., Chan, E., Davison, D., Earls, M., Evans, S. W., Flinn, S. K., Froehlich, T., Frost, J., Holbrook, J. R., Lehmann, C. U., Lessin, H. R., Okechukwu, K., Pierce, K. L., Winner, J. D., & Zurhellen, W. (2019). Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics (Evanston)*, *144*(4). https://doi.org/10.1542/peds.2019-2528

Wong, A.W.W.A., & Landes, S. D. (2022). Expanding understanding of racial-ethnic differences in ADHD prevalence rates among children to include Asians and Alaskan Natives/American Indians. *Journal of Attention Disorders*, *26*(5), 747–754. https://doi.org/10.1177/10870547211027932

Xu, J., Murphy, S. L., Kochanek, K. D., & Arias, E. (2022). Mortality in the United States, 2021. *NCHS Data Brief*, *456*, 1–8. https://dx.doi.org/10.15620/cdc:122516

Yang, K. G., Flores, M. W., Carson, N. J., & Cook, B. L. (2022). Ethnic disparities in childhood ADHD treatment access and utilization: Results from a national study. *Psychiatric Services (Washington, D.C.)*, *73*(12), 1338–1345. https://doi.org/10.1176/appi.ps.202100578

Yang, Y., Harkness, J. A., Chin, T., & Villar, A. (2010). Response styles and culture. In *Survey Methods in Multinational, Multiregional, and Multicultural Contexts* (pp. 203–223). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470609927.ch12

Tables and Figures

Table 1

SDQ-HI Items

SDQ Item	Item Description
1	Sees tasks through to the end, good attention span
2	Easily distracted, concentration wanders
3	Constantly fidgeting or squirming
4	Restless, overactive, cannot stay still for long
5	Thinks things out before acting

Table 2

Descriptive Statistics (N = 6,673)

Variables	*n*	%	*M*	*SD*
Age			11.1	4.06
Sex				
Male	3,460	51.9		
Female	3,212	48.1		
Missing	1	0		
Ethnicity				
Non-Hispanic White	3,779	55.6		
Hispanic/Latinx	1,670	25		
Non-Hispanic Black	828	12.4		
Non-Hispanic Asian	396	5.9		
Child's Relationship with Adult Informant				
Parent (biological, adoptive, or step)	6,240	93.5		
Grandparent	281	4.2		
Brother/sister	52	0.8		
Other relative/non-relative	51	0.8		
Aunt/uncle	40	0.6		
Missing	9	0.1		
Nativity Status				
U.S. born	6,323	94.8		
Foreign born	326	4.9		
Missing	24	0.4		
Highest Level of Education in the Family				
B.A. degree	1,691	25.3		
High school graduate or G.E.D.	1,155	17.3		
M.A. degree	1,069	16		
A.A. degree	1,062	15.9		
Some college, no degree	959	14.4		
Professional or doctoral degree	382	5.7		
Grades 1-12, no diploma	350	5.2		
Missing	5	0.1		
Total Combined Family Income				
$100,000 and over	2,503	37.5		
$0-34,999	1,381	20.7		
$50,000-74,999	1,142	17.1		
$75,000-99,999	877	13.1		
$35,000-49,999	770	11.5		
ADHD Diagnosis				
Ever Diagnosed	685	10.3		
Never Diagnosed	5,972	89.5		
Missing	16	0.2		
SDQ-HI			2.17	2.45

Table 3

Model Fit Assessment for SDQ-HI (N = 6,673)

Model	χ^2	*df*	RMSEA (90% CI)	SRMR	CFI	TLI
Model 1	446.980^{***}	5	.115 (.106-.124)	.045	.967	.934
Model 2	375.355^{***}	4	.118 (.108-.128)	.042	.972	.930
Model 3	291.850^{***}	4	.104 (.094-.114)	.032	.978	.946
Model 4	70.401^{***}	3	.058 (.047-.070)	.017	.995	.983
χ^2 Difference Tests	$\Delta\chi^2$	Δdf				
Model 1 vs. Model 2	79.767^{***}	1				
Model 3 vs. Model 4	184.651^{***}	1				

Note. $^{***}p < .001$

Table 4

Model Fit Assessment for SDQ-HI Across Ethnic Groups

Ethnic Group	Model	$\chi^2/\Delta\chi^2$	$df/\Delta df$	RMSEA (90% CI)	SRMR	CFI	TLI
White	Model 3	300.842^{***}	4	.140 (.127-.154)	.032	.982	.956
(n = 3,779)	Model 4	87.847^{***}	3	.087 (.071-.103)	.019	.995	.983
	Model 3 vs. Model 4	180.663^{***}	1				
Hispanic/ Latinx	Model 3	91.722^{***}	4	.115 (.095-.136)	.035	.974	.936
	Model 4	25.777^{***}	3	.067 (.045-.093)	.019	.993	.978
(n = 1,670)	Model 3 vs. Model 4	61.060^{***}	1				
Black	Model 3	68.589^{***}	4	.140 (.112-.170)	.040	.976	.941
(n = 828)	Model 4	17.383^{***}	3	.076 (.044-.112)	.020	.995	.982
	Model 3 vs. Model 4	41.272^{***}	1				
Asian	Model 3	15.397^{**}	4	.085 (.043-.132)	.043	.979	.947
(n = 396)	Model 4	8.072^{*}	3	.065 (.009-.969)	.030	.991	.969
	Model 3 vs. Model 4	6.574^{*}	1				

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 5

Model Fit Assessment for Configural Invariance Model (N = 6,673)

	χ^2	*df*	RMSEA (90% CI)	SRMR	CFI	TLI
Configural Invariance	134.770^{***}	12	.078 (.067-.090)	.020	.995	.983

Note. $^{***}p < .001$

Table 6

Standardized Factor Loadings and Item Thresholds (N = 6,673)

Item	Factor Loadings (λ)					Threshold	Item Thresholds (τ)				
	λ_{White}	$\lambda_{Hispanic}$	λ_{Black}	λ_{Asian}	R^2		τ_{White}	$\tau_{Hispanic}$	τ_{Black}	τ_{Asian}	R^2
SDQ1	.778	.698	.721	.726	.925	SDQ1$1	.501	.702	.588	.690	.869
						SDQ1$2	2.049	1.934	1.899	2.076	.903
SDQ2	.958	.937	.947	.892	.521	SDQ2$1	1.107	1.239	1.354	.755	0
						SDQ2$2	4.118	3.553	3.792	2.900	0
SDQ3	.912	.805	.930	.622	.691	SDQ3$1	1.721	1.037	2.052	1.282	.108
						SDQ3$2	3.276	2.242	3.529	**2.102**	0
SDQ4	.888	.819	.886	.772	.999	SDQ4$1	1.242	1.228	1.112	1.127	.749
						SDQ4$2	2.935	**2.447**	2.548	2.412	0
SDQ5	.595	.568	.554	.627	0	SDQ5$1	-.077	**.171**	-.137	.182	.290
						SDQ5$2	1.572	1.651	1.353	1.798	.269

Note. Bolded parameters indicate non-invariance at $p < .05$.

Table 7

Model Fit Assessment for Configural Invariance Models for Pairwise Comparisons

	χ^2	*df*	RMSEA (90% CI)	SRMR	CFI	TLI
White vs. Hispanic/Latinx	117.025^{***}	6	.082 (.070-.096)	.019	.995	.982
White vs. Black	102.438^{***}	6	.084 (.070-.098)	.019	.995	.984
White vs. Asian	81.348^{***}	6	.078 (.063-.093)	.020	.996	.987

Note. $^{***}p < .001$

Table 8

Standardized Factor Loadings and Item Thresholds

	Factor Loadings (λ)				Item Thresholds (τ)			
	Item	λ_{White}	$\lambda_{Hispanic}$	R^2	Threshold	τ_{White}	$\tau_{Hispanic}$	R^2
White vs. Hispanic/ Latinx	SDQ1	.778	.700	.997	SDQ1$1	**.315**	**.532**	-
					SDQ1$2	1.287	1.414	.810
	SDQ2	.957	.931	.998	SDQ2$1	.318	.471	0
					SDQ2$2	1.182	1.280	.502
	SDQ3	.912	.804	.849	SDQ3$1	.707	.807	.373
					SDQ3$2	1.346	1.469	.199
	SDQ4	.888	.820	1	SDQ4$1	.570	.736	.106
					SDQ4$2	**1.348**	**1.435**	-
	SDQ5	.595	.568	0	SDQ5$1	**-.062**	**.163**	-
					SDQ5$2	1.263	1.381	0
	Item	λ_{White}	λ_{Black}	R^2	Threshold	τ_{White}	τ_{Black}	R^2
White vs. Black	SDQ1	.778	.723	.996	SDQ1$1	.315	.438	0
					SDQ1$2	1.287	1.344	.807
	SDQ2	.959	.942	.997	SDQ2$1	.318	.474	0
					SDQ2$2	1.183	1.259	.704
	SDQ3	.912	.927	0	SDQ3$1	.707	.819	.387
					SDQ3$2	1.346	1.364	.460
	SDQ4	.888	.888	.987	SDQ4$1	.570	.577	.993
					SDQ4$2	1.347	1.246	0
	SDQ5	.595	.554	.063	SDQ5$1	-.062	-.076	.481
					SDQ5$2	1.264	1.165	0
	Item	λ_{White}	λ_{Asian}	R^2	Threshold	τ_{White}	τ_{Asian}	R^2
White vs. Asian	SDQ1	.778	.723	.732	SDQ1$1	.315	.489	.848
					SDQ1$2	1.287	1.441	.984
	SDQ2	.958	.898	.772	SDQ2$1	.318	.354	.043
					SDQ2$2	1.182	1.326	0
	SDQ3	.912	.623	.704	SDQ3$1	.707	1.053	0
					SDQ3$2	**1.346**	**1.697**	-
	SDQ4	.888	.771	1	SDQ4$1	.570	.780	.995
					SDQ4$2	1.348	1.596	0
	SDQ5	.595	.626	0	SDQ5$1	-.062	.192	.490
					SDQ5$2	1.264	1.451	.519

Note. Bolded parameters indicate non-invariance at $p < .05$.

Table 9

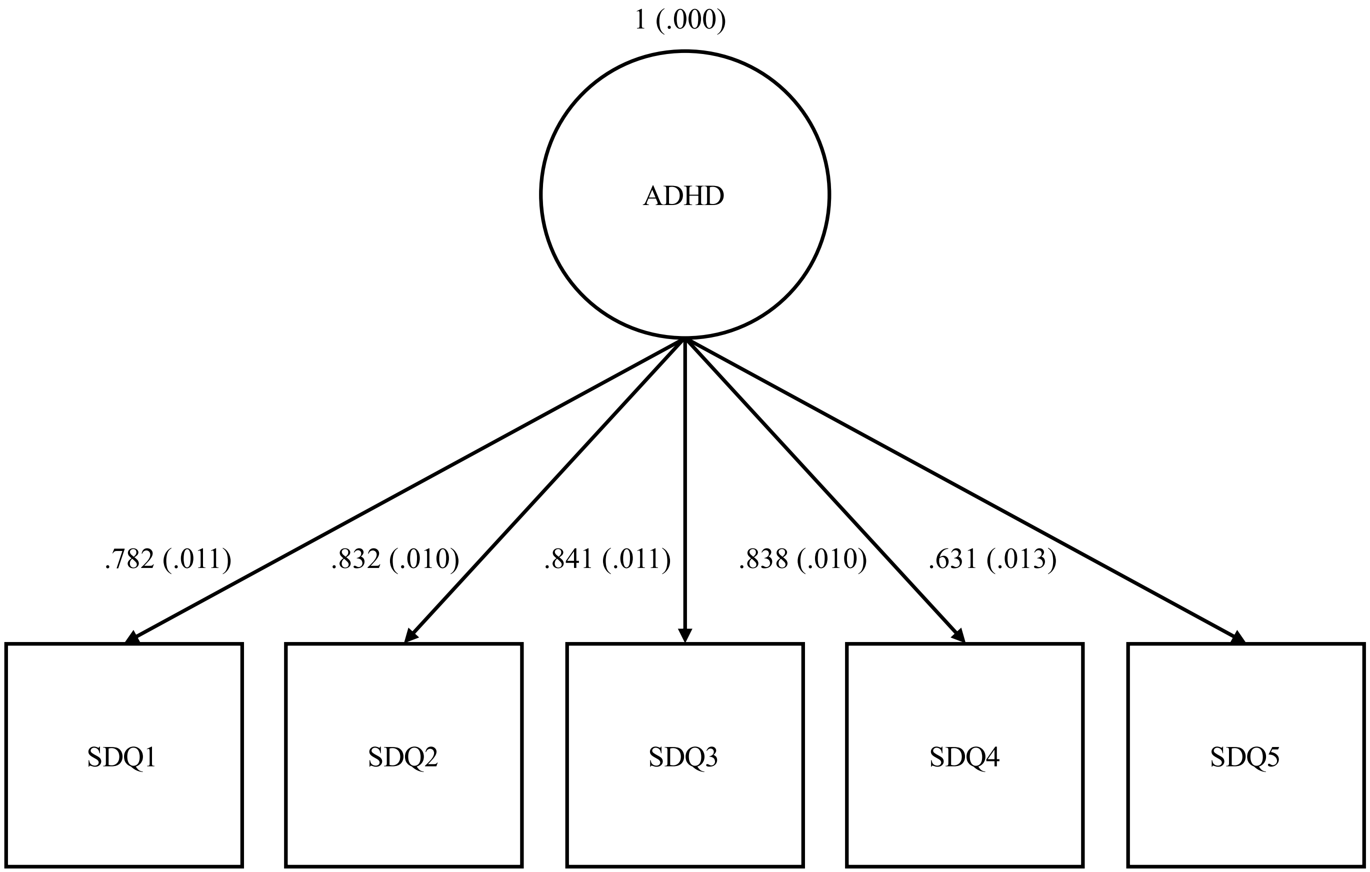
Pairwise Comparisons of Latent Means

	Hispanic/Latinx[a]	Black[b]	Asian[c]
Inattention	.093	.127	-.347
Hyperactivity-Impulsivity	.062	-.055	-.228

Note. Each ethnic group is compared with the White reference group.
a. In the White-Hispanic/Latinx comparison, the Hispanic/Latinx group ranked higher on inattention and hyperactivity-impulsivity.
b. In the White-Black comparison, the Black group ranked higher on inattention and lower on hyperactivity-impulsivity.
c. In the White-Asian comparison, the Asian group ranked lower on inattention and hyperactivity-impulsivity.

Figure 1

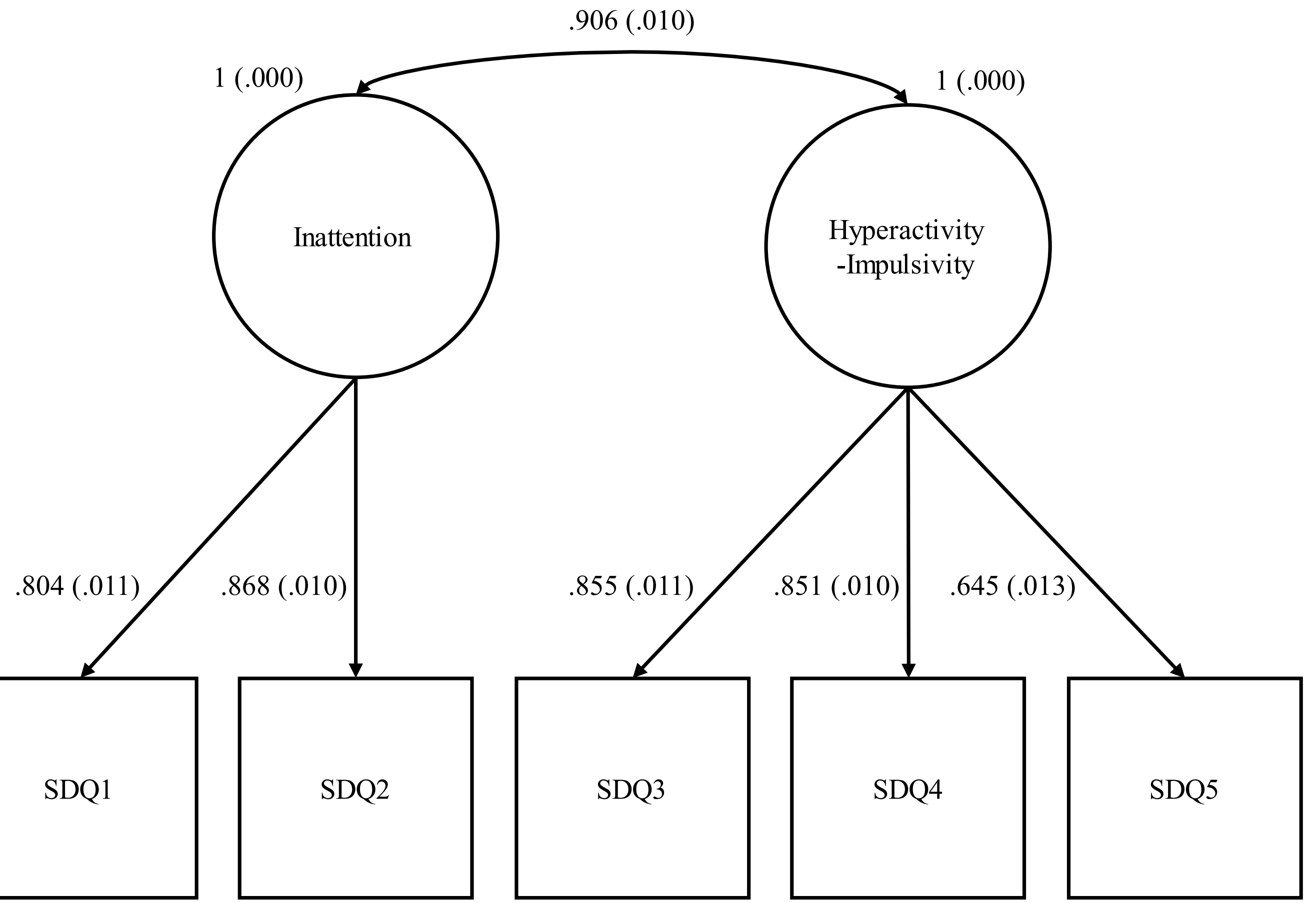
One-Factor SDQ-HI Model



Note. Standardized parameter estimates are presented along with standard errors in parentheses. All estimates are significant at $p < .001$.

Figure 2

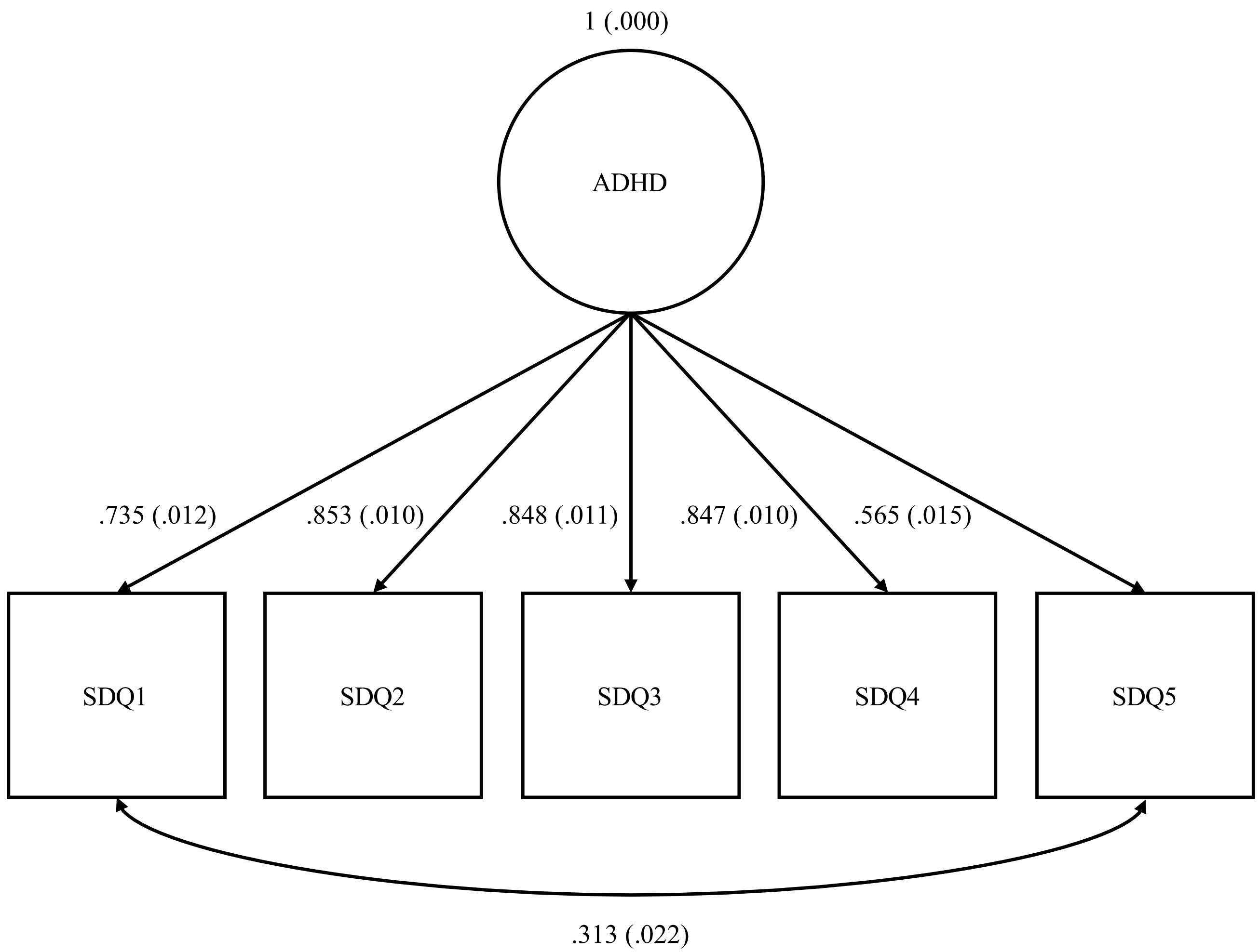
Two-Factor SDQ-HI Model



Note. Standardized parameter estimates are presented along with standard errors in parentheses. All estimates are significant at $p <$ .001.

Figure 3

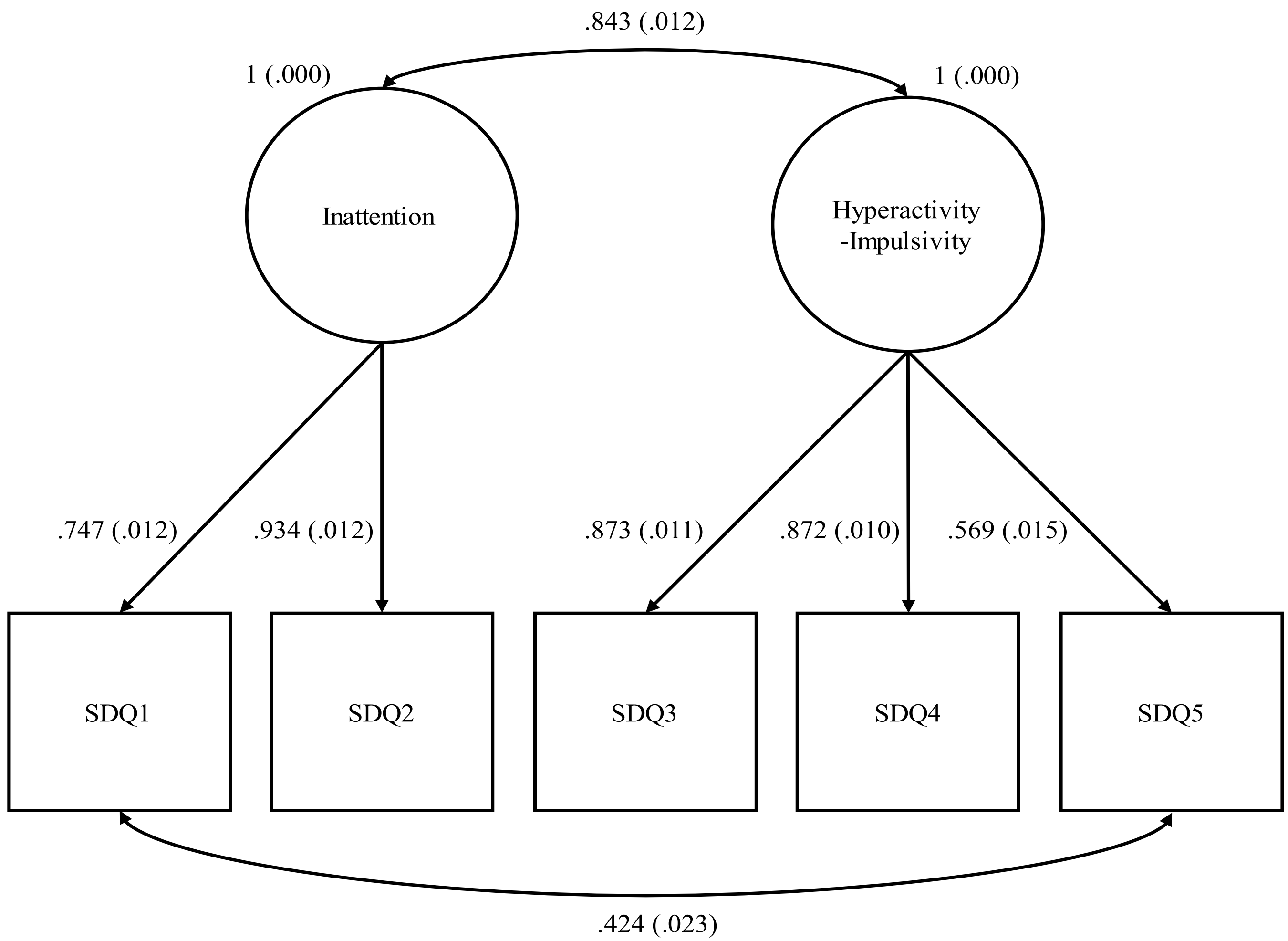
Modified One-Factor Model



Note. Standardized parameter estimates are presented along with standard errors in parentheses. All estimates are significant at $p <$ .001.

Figure 4

Modified Two-Factor Model



Note. Standardized parameter estimates are presented along with standard errors in parentheses. All estimates are significant at $p <$ .001.

Appendix A: Strengths and Difficulties Questionnaire

Strengths and Difficulties Questionnaire

P or T 4-10

For each item, please mark the box for Not True, Somewhat True or Certainly True. It would help us if you answered all items as best you can even if you are not absolutely certain. Please give your answers on the basis of the child's behavior over the last six months or this school year.

Child's name .... Male/Female

Date of birth..

	Not True	Somewhat True	Certainly True
Considerate of other people's feelings	☐	☐	☐
Restless, overactive, cannot stay still for long	☐	☐	☐
Often complains of headaches, stomach-aches or sickness	☐	☐	☐
Shares readily with other children, for example toys, treats, pencils	☐	☐	☐
Often loses temper	☐	☐	☐
Rather solitary, prefers to play alone	☐	☐	☐
Generally well behaved, usually does what adults request	☐	☐	☐
Many worries or often seems worried	☐	☐	☐
Helpful if someone is hurt, upset or feeling ill	☐	☐	☐
Constantly fidgeting or squirming	☐	☐	☐
Has at least one good friend	☐	☐	☐
Often fights with other children or bullies them	☐	☐	☐
Often unhappy, depressed or tearful	☐	☐	☐
Generally liked by other children	☐	☐	☐
Easily distracted, concentration wanders	☐	☐	☐
Nervous or clingy in new situations, easily loses confidence	☐	☐	☐
Kind to younger children	☐	☐	☐
Often lies or cheats	☐	☐	☐
Picked on or bullied by other children	☐	☐	☐
Often offers to help others (parents, teachers, other children)	☐	☐	☐
Thinks things out before acting	☐	☐	☐
Steals from home, school or elsewhere	☐	☐	☐
Gets along better with adults than with other children	☐	☐	☐
Many fears, easily scared	☐	☐	☐
Good attention span, sees work through to the end	☐	☐	☐

Signature .... Date ....

Parent / Teacher / Other (Please specify):

Thank you very much for your help