THE VALIDITY OF ADHD BEHAVIORAL RATING SCALES ACROSS ETHNIC GROUPS

by

ELIZABETH ANN CURKO KERA

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy,

The City University of New York

2006

UMI Number: 3232013

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted by the Graduate Faculty in Psychology in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

_____	_____
Date	Jeffrey M. Halperin, Ph.D. Chair of Examining Committee
_____	_____
Date	Joseph Glick, Ph.D. Executive Officer

Howard Ehrlichman, Ph.D.

Tina Moreau-Jones, Ph.D.

Murphy Halliburton, Ph.D.

Carol Friedman, Ph.D.

Supervisory Committee

The City University of New York

ABSTRACT

THE VALIDITY OF ADHD BEHAVIORAL RATING SCALES ACROSS ETHNICITIES

by

ELIZABETH ANN CURKO KERA

Advisor: Jeffrey M. Halperin, Ph.D.

Objective: Behavioral rating scales are commonly used to assess symptomatology related to Attention-Deficit/Hyperactivity Disorder (ADHD). However, parent and teacher ratings vary considerably as a function of ethnicity. Thus, it is unclear whether there is differential validity of behavioral rating scales across ethnic groups or whether there are real differences in the behavior of children of different ethnicities. This study used objective laboratory measures to validate parent- and teacher-rated ADHD symptoms as assessed by the ADHD-IV Rating Scale in four different ethnicities: European American, African American, Asian, and Hispanic. It was hypothesized that ethnic groups would differ on subjective ratings, but not on objective measures. **Methods:** Participants were European American (N=50), African American (N=13), Asian (N=13), and Hispanic (N=10) preschoolers who were rated as either having a high or low number of symptoms by parent and teacher report. Total number of errors on a children's continuous performance task and solid-state actigraph data were used to validate subjective ratings of ADHD symptoms. **Results:** Teachers rated African

American children significantly higher than European American children ($p<.01$). A similar trend was found to for parent ratings between those two groups ($p=.056$). Significant correlations were found for parent and teacher agreement of ADHD symptomatology in the European American, Asian, and Hispanic group; however, poor agreement was found for the African American group. The Hispanic group demonstrated the best agreement between subjective and objective measures, while the European American group demonstrated adequate validity. However, agreement of subjective and objective measures was virtually non-existent in the African American and Asian groups. **Conclusions:** Although findings indicate generally adequate validity of ADHD behavioral rating scales in Hispanic and European American children, the ratings do not appear to be valid for African American and Asian children. Thus, both parent and teacher rating scales should be interpreted with caution when used in these groups.

ACKNOWLEDGEMENTS

My interest in the field of neuropsychology originally emerged in high school, after becoming fascinated with the brain-behavior relationship. This interest grew even stronger over the course of my undergraduate studies. I made the decision to pursue my Ph.D. in the field, and shortly thereafter upon entering graduate school began working with Dr. Jeffrey Halperin. I can truly say that I consider myself lucky to have found such a wonderful and dedicated mentor who encourages his students to find something they are passionate about in their career and pursue it. Through his mentorship, support, and guidance, I was able to develop and strengthen my skills as a developmental neuropsychologist in both clinical and research settings. I am truly grateful for all of the help he has provided me throughout my years as a graduate student, and particularly look forward to long-lasting future collaborations. I would also like to thank my dissertation committee members, Dr. Howard Ehrlichman, Dr. Tina Moreau-Jones, and Dr. Murphy Halliburton, as well as my outside reader, Dr. Carol Friedman for all of their hard work, support, and encouragement. I could not have asked for a better dissertation committee- as they not only added to the integrity of my study, but were invaluable during the dissertation process.

This study would not have been possible were it not for the members of the Queens College Preschool Project Team: Katia Policaro, Dr. David Marks, Olga Berwid, Amita Santra, Olga Magidina, Chaya Gopin, Taneka Wellington, Michelle Bubnik, Bella Hahn, Dr. Carlin Miller, and Dr. Dione Healey. I am thankful not only for their hard work, but for their support and friendships as well.

For all those who listened to the moans and groans along the way, but miraculously were able to get me to re-focus on the fact that there IS a light at the end of the tunnel- Sandy, Peter, Tina, Olga, as well as many other family members and good friends- I thank you for simply being there to listen, as well as for your encouragement.

To my husband Ivan, for all of his love and support, I am forever grateful. We vowed for better or for worse- but helping me enter references into Reference Manager was not part of the deal- and he did it anyway. Now that's what I call a catch. I love you dragi.

Finally, I would like to thank my parents, without whom I quite literally would not be who or where I am today. They came to this country from Croatia with almost nothing, but were able to give their children just about everything. They made it so that I had no excuse to not pursue my dreams or fulfill my potential. I can only hope to be able to give my children the same opportunities. In my eyes, the amount of effort and dedication I have put into getting this degree pales in comparison to the hard work and sacrifice they have endured to get me to this point. It is their example that got me through the many moments of wondering why I put myself through this grueling process. They have influenced me more than they will ever know, and I share this accomplishment with them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SPECIFIC AIMS

Attention Deficit/Hyperactivity Disorder (ADHD) is a developmental disorder marked by symptoms of inattention, hyperactivity, and impulsivity that usually emerge during the preschool years. Parent and teacher behavioral rating scales are often used to identify children with ADHD symptoms and play a central role in the diagnostic process. However, the use of such subjective measures could lend itself to various biases and cultural variations, which may have implications for their validity across ethnic groups.

Several studies have shown differences in ratings of children as a function of ethnicity (Baumgaertel, Wolraich, & Dietrich, 1995; Dominguez de Ramirez & Shapiro, 1998; Graetz, Sawyer, Hazell, Arney, & Baghurst, 2001; Leung et al., 1996; Magnusson, Smari, Gretarsdottir, & Prandardottir, 1999; O'Leary, Vivian, & Nisi, 1985a; Pineda, Lopera, Palacio, Ramirez, & Henao, 2003; Taylor & Sandberg, 1984). These differences in reported prevalence of ADHD symptoms across ethnic groups may be due to culture-bound biases and/or assumptions, such that the ratings of the child differ as a function of ethnicity of either the rater (i.e., parent or teacher) or the child (Dominguez de Ramirez et al., 1998; Dominguez de Ramirez, 2001). Yet, the extent to which these differences in ratings reflect true differences in the child's behavior vs. differences in the percepts of the child as a function of rater or child ethnicity remains unclear. In view of the increasing ethnic heterogeneity in the United States and the important role that behavioral ratings play in the diagnosis and treatment of ADHD, it is critical that such ratings be assessed for validity across ethnic groups to determine whether they are assessing the same

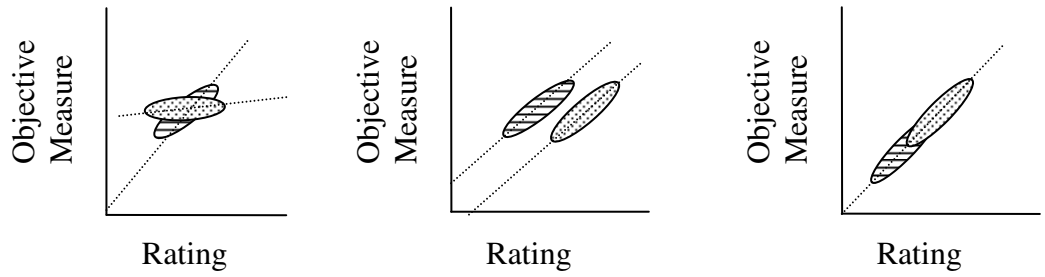
construct, and/or whether there is a need for different norms or clinical cut points for different ethnic groups.

As is the case with most psychological measures, there is no true "gold standard" for validating ratings of ADHD behaviors by parents and teachers. Therefore, one typically evaluates construct validity by examining the relationship between the target measure (i.e., ratings) and multiple measures of independently-derived yet theoretically-related constructs. While a common method of validating ratings is using the clinician's judgment based on an interview with the parent or teacher, this approach is not independent of the initial ratings. Additionally, the fact that these two individuals are oftentimes in close contact further confounds the independence of such reports. As such, reliability along with face validity, rather than true validity, oftentimes becomes the primary standard by which child psychopathology rating scales are judged. Nevertheless, if systematic biases as a function of ethnicity are present, ratings may be "reliably wrong."

Another approach to assessing the validity of scales across ethnic groups is to use objective laboratory-based measures of ADHD symptom domains to validate parent and teacher ratings. While few would argue that objective measures should be used for diagnostic purposes, and correlations between ratings and objective measures are not robust, they are fairly consistently found to be statistically significant (Barkley, 1991; Halperin, Matier, Bedi, Sharma, & Newcorn, 1992). Thus, the use of objective data to validate parent and teacher ratings can provide specific information as to whether ratings have similar or differential validity across ethnic groups. Further, if differential validity across ethnic groups is found, it is important to determine the nature of such differences.

Figure 1 depicts three possible outcome scenarios, with the objective criterion plotted on the y-axis and the subjective behavioral rating on the x-axis. If differences are due to differential validity coefficients, or "slope bias," such that significant correlations are found in only some groups, this would indicate that the scales are inappropriate for use with other groups (see Figure 1a). In contrast, if regression lines (or correlations) across ethnic groups are similar, but have differential Y-intercepts (i.e., "intercept bias"), as would occur if groups differ on ratings but not objective measures, this would likely indicate validity across groups but the need for distinct ethnically-specific cutoffs (see Figure 1b). Finally, if groups differ in the same direction on both ratings and objective measures, and a single regression line characterizes the relationship in both groups, this would suggest true differences in the manifestation of ADHD as a function of ethnicity (see Figure 1c).

Figure 1a. **Figure 1b.** **Figure 1c.**



Note: Ethnicity A and B are differentiated by stripes and dots.

This study's goal was to evaluate the appropriateness of behavioral rating scales across various ethnic groups. It was hypothesized that:

1) Parents and teachers would rate children differently as a function of ethnicity such that, African American children would be rated higher than European American children, who in turn would be rated higher than Asian children. It was hypothesized that Hispanic[1] children would be rated similarly to European American children, regardless of whether they are white or black Hispanic. In contrast to parent and teacher ratings, objective data would not differ as a function of ethnicity.[2]

2) Slope bias would not exist; the scales themselves would be valid and useful in assessing the same construct across ethnic groups. However, intercept bias would emerge such that different cut-off points would need to be determined for the use of appropriate norms for children based on ethnic group.

[1] In this manuscript, children who meet inclusion criteria for the Hispanic group (see Methods) were called "Hispanic"; however, the reader should note that in previous studies reviewed in this manuscript, children of similar ethnic backgrounds may have been included in a group called "Hispanic" or "Latino." Although inclusion criteria varied across studies, similar ethnic backgrounds were generally included into these "Hispanic" and "Latino" groups. Where applicable, the author used the original studies' name for each subject group. Therefore the reader should keep in mind that these groups may in fact similar and comparable, although there may be different names used.

[2] It should be noted that children of Asian and Hispanic descent could also be called Asian-American and Hispanic-American. Being that those particular terms are not commonly used to refer to those groups as they are in the African American and European American communities, in this study those groups were referred to as Asian and Hispanic.

INTRODUCTION

Overview of ADHD

Attention-Deficit/Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder characterized by symptoms of inattention, hyperactivity, and impulsivity that are not age-appropriate. Although the disorder has received attention primarily in the past few decades, ADHD was in fact first described in medical journals at the beginning of the last century (Still, 1902) when a British physician described a group of children as having "a defect in moral control" (p. 1009). He indicated that these children had characteristics of impulsivity, aggression, overactivity, and deficits in inhibition; characteristics that according to modern diagnostic criteria would describe children with ADHD. The validity of the disorder has been questioned in recent years due not only to negative attention in the popular media, but also to doubt in the scientific community surrounding its boundaries and differentiation from other psychiatric conditions (e.g., Oppositional Defiant Disorder, Conduct Disorder). Nevertheless, abundant data certainly support the existence of a valid and very impairing disorder for those afflicted. At a recent international consortium (Barkley, Cook, & et al., 2002) reviewing scientific literature regarding the etiology, nature, diagnostic criteria, and interventions for ADHD, leading researchers in the field concluded that "… as a matter of science, the notion that ADHD does not exist is simply wrong. All of the major medical associations and government health agencies recognize ADHD as a genuine disorder because the scientific evidence indicating it is so overwhelming" (p. 89). Furthermore, it was indicated that "… inaccurate stories rendering ADHD as myth, fraud, or benign condition may cause

thousands of sufferers not to seek treatment for their disorder…" (p. 89), a conclusion which emphasizes the dangerous impact that such negative attention may have on the outcome of the disorder in many patients.

The technical name for the disorder has changed over the years to reflect not only the growing knowledge, but also etiological theory of the disorder. Terms such as Minimal Brain Dysfunction (MBD), Hyperkinetic Impulsive Disorder, and Attention Deficit Disorder (ADD) have all been used to describe the same or similar collection of symptoms and behaviors that we currently recognize as constituting Attention-Deficit/Hyperactivity Disorder. The current DSM-IV diagnostic structure of the disorder describes ADHD as being comprised of deficits in two main domains: inattention and hyperactivity/impulsivity. Based on this two-factor structure, an affected individual can be diagnosed with one of three main subtypes of ADHD: 1) ADHD-inattentive type (ADHD-I), 2) ADHD-hyperactive/impulsive type (ADHD-HI), and 3) ADHD-combined type (ADHD-C). Therefore, the type of ADHD one has simply describes which symptoms of ADHD are most prominent and impairing to the individual, those primarily of inattention, hyperactivity/impulsivity, or both.

Today, ADHD is among the most commonly diagnosed disorders of childhood with prevalence estimates between 3-7% (AmericanPsychiatricAssoc., 1994). The symptoms usually emerge during the early childhood years and clinically significant impairment from these symptoms in more than one setting (e.g., home, school, work) is required to meet current DSM-IV diagnostic criteria. One must also have symptoms present before age seven, and the symptoms must persist for at least six months.

Furthermore, the symptoms of ADHD cannot be better accounted for by another disorder (e.g., inattention due to depression).

Although once thought to disappear after childhood, symptoms of ADHD can in fact be present throughout the lifespan (Shaffer, 1994). Furthermore, the disorder is developmental in nature, such that the symptomatology typically changes over the lifespan. While symptoms of hyperactivity and impulsivity are commonly found in young children, these individuals increasingly evidence symptoms of inattention, as opposed to overactivity, as they mature into adulthood (Wilens, Biederman, & Spencer, 2002). Symptoms of hyperactivity/impulsivity and inattention in affected individuals can be quite impairing, not only to themselves, but also to surrounding family, friends, and society. Individuals with ADHD have been shown to have higher rates of other psychopathology (e.g., conduct, substance abuse, and anxiety disorders), criminal offenses, school drop-out, and difficulty sustaining employment (Biederman, Newcorn, & Sprich, 1991; Hechtman & Weiss, 1986; Millstein, Wilens, Biederman, & et al., 1997). In addition, families in which the child has been diagnosed with ADHD have been reported to have higher rates of family dysfunction, marital discord, parent psychopathology, and parental stress (Johnston & Mash, 2001). However, it is not yet clear the degree to which this adverse family life is due to the child's psychopathology, or other variables such as pre-existing parent psychopathology, socioeconomic status, or all of the above in combination. It is now well-established that ADHD has a substantial genetic component (Todd et al., 2005; Madras, Miller, & Fischman, 2005; Faraone et al., 2005), and as such, parents of children with ADHD are also frequently symptomatic.

Depending on factors such as the nature and severity of symptoms, as well as the age of the individual, common treatments for ADHD can include interventions taken from a variety of different treatment modalities. Typical treatment options for children include behavior modification programs, parent skills training, and medication treatment, while typical treatment options for adults typically involve cognitive-behavioral based individual or group therapy, and medication treatment. Multi-modal treatment (i.e., a combination of behavioral therapy and medication) is typically accepted as the most effective intervention (Barkley, 2000). However, data supporting this assertion have been somewhat inconsistent, as some studies have found that medication alone is just as effective as medication and behavioral therapy combined (MTA Cooperative Group, 1999). Within the multi-modal treatment approach, symptoms of ADHD (e.g., hyperactivity, inattention) are treated with medication, therefore giving the patient the opportunity to work on improving other areas of difficulty (e.g., social skills, learning problems, self-control) which may be less improved by the medication. Of note, there has been growing concern as to whether children with ADHD are being over-identified, and as a result, over-medicated for their symptoms. Many health care professionals are increasingly reluctant to prescribe medication in younger children unless absolutely necessary (i.e., symptoms are very severely impairing) as it is still not clear what possible long-term consequences medication treatment may have on brain development. Given the fact that approximately half of all preschoolers that appear to exhibit ADHD-like symptoms grow out of these behaviors (Campbell, 2002), it is imperative that valid assessment methods (e.g., behavioral rating scales) be used during the diagnostic process to gather information related to the evaluation of ADHD symptoms. This would ensure

that the health care provider is not only gathering accurate information about the possible existence of the disorder in the child, but would also aid in the decision of whether to pursue various treatment methods.

ADHD assessment procedures: The use of rating scales

The most reliable diagnosis of ADHD involves a multimethod approach in which information from parents, teachers, and the child are considered together (DuPaul, Guevremont, & Barkley, 1991; Mitsis, McKay, Schulz, Newcorn, & Halperin, 2000; Power et al., 1998; Power, Costigan, Leff, Eiraldi, & Landau, 2001). Gathering information from the child typically involves administering a test battery which can include standardized tasks such as intelligence and neuropsychological measures, as well as clinical observation of the child. These methods of assessment typically provide the clinician with important information as to how the child is currently functioning in areas of cognition, academics, and social skills. However, it is well known that within the context of a one-on-one structured setting, children with ADHD oftentimes appear attentive and normo-active, thus limiting the diagnostic utility of clinical observations in such a setting. Further, individuals, and particularly children, with ADHD have limited insight into their own impairment, thus limiting the clinical utility of self-reports.

As a result, gathering information from parents and teachers is essential for the evaluation of ADHD in children, and to facilitate the collection of information from these informants in a standardized manner, ratings have become a mainstay of the field. The DSM-IV as well as the ICD-10 require impairment due to ADHD symptoms to be evident in more than one setting for the diagnosis of the disorder, and for young children,

the home and school settings are most commonly evaluated. Oftentimes, a clinician will administer a clinical interview to the parent, but not get the same opportunity to gather such detailed information and descriptions from the teacher. Furthermore, home and/or school observations by the clinician are rarely possible. Thus, the most typical means of assessing the cross-situational presence of ADHD symptoms is via the use of parent and teacher behavioral rating scales. Such scales have become universally accepted tools to assist in the diagnosis of ADHD. However, assessing the validity of these scales can be somewhat difficult, as there currently is no "gold standard" against which to validate these rather subjective ratings. The most commonly used method of validating such ratings is by using a clinician's judgment based on parent/teacher interview; however, this information is not independent of the initial ratings. Moreover, information obtained from parents and teachers may not be independent of each other, as these two individuals are often in close contact with one another. Lastly, information gathered about the child's behavior in more than one setting can be obtained solely from one informant (such as the parent), and therefore the ratings of the child across settings are not independent of one another.

Of interest to the evaluation of cross-situational presence of symptoms is the degree to which parents and teachers agree on whether symptoms in the child exist, and if so, to what extent. Concordance rates of parent- and teacher-reported ADHD symptoms, although generally found to be statistically significant, tend to be only modest (Wolraich et al., 2004; Mitsis et al., 2000; Antrop, Roeyers, Oosterlaan, & Van Oost, 2002), with reported correlations ranging from .25 to .45. This limited agreement may be due to several factors. First, the parent and teacher see the child in very different settings. While

the parent often sees the child in his/her family environment, the teacher sees the child in a more structured environment where there are typically many more cognitive and behavioral demands placed on the child. Second, teachers generally have a more diverse base for comparison, i.e., they can compare the child's behavior against a larger comparison group because they work with many children throughout the day. Therefore, teachers presumably have a more accurate sense of whether the child's behavior deviates from the expected age-appropriate norm. Parents on the other hand, may only have experience with that child, or may be limited to comparing their child's behavior against a very small group of children which would likely include siblings or cousins. The role that common environmental influences had in shaping the children's behavior may be a confound, and therefore provide for an unrepresentative comparison group. Lastly, children might behave quite differently in such varied environments such as school and home for a variety of reasons which can include different environmental structure or stimulation, parent vs. teacher relationship with the child, varied disciplinary styles, and behavior in front of peers. Due to this variability, the child may be likely to exhibit different behaviors in the presence of different caregivers, and therefore these differences would likely be emphasized in parent and teacher ratings of the same child.

In other ways, however, teacher ratings may be more prone to bias that is not as apparent in parent ratings, such as a possible negative impact of other children in the classroom on the child's behavior. For example, data have supported the idea of teacher-rated halo effects, such that ADHD symptoms are significantly inflated when comorbid with other disruptive behavior symptomatology (e.g., ODD symptoms) (Abikoff, Courtney, Pelham, & Koplewicz, 1993; Abikoff et al., 2002). These findings have been

found to be unidirectional in nature (i.e., ADHD ratings are inflated when ODD symptoms are present, but ratings of ODD are not inflated in the presence of ADHD symptoms), and is thus especially critical in the interpretation of ADHD behavioral ratings. In addition, teacher ratings may in some cases be more susceptible to being biased as a function of ethnicity because there is more often a mismatch between teacher and child ethnicity than parent and child ethnicity. Therefore, differences in ratings across ethnic groups may or may not be indicative of "real" differences in the child's behavior. Instead they may be a reflection of other factors that may include such things as parent/teacher perceptions of the child, the rater's mood on that particular day, as well as cultural expectations based on either the child's or the rater's ethnicity.

Psychometric properties of commonly-used rating scales

As noted earlier, due to DSM-IV and ICD-10 requirements that symptoms be cross-situational (i.e., present in two or more settings), behavioral scales have become a mainstay in the diagnostic process, especially since they are among the most convenient ways of gathering information from multiple informants. There are many behavioral rating scales which assess ADHD symptoms that are currently used during the diagnostic evaluation process in conjunction with clinical interview, observational, and test data. These scales are designed to assess for the presence of ADHD symptoms, and in some cases, symptoms related to other psychopathology as well. Given the fact that behavioral ratings have become such an important tool in the evaluation process, it would be important to review the known psychometric properties (i.e., validity, reliability, characteristics of normative group) of scales relevant to this study. Especially important

to this study would be whether the authors who designed the scales initially found any differences in ratings as a function of ethnic group in the samples on which the rating scale data were normed.

Scales such as the ADHD-IV Rating Scale (DuPaul, Power, Anastopoulos, & Reid, 1998) assess the basic ADHD symptoms by listing the eighteen DSM-IV criteria, then requiring the parent or teacher to rate the frequency of each behavior on a 4-point Likert scale. Other rating scales such as the SNAP-IV Teacher and Parent Rating Scales (Swanson & Carlson, 1994) are somewhat more expansive and assess not only for ADHD symptomatology but also symptoms of other common comorbid disorders as well, such as Oppositional Defiant Disorder. Lastly, some behavioral rating scales are lengthier, and assess an even broader range of child psychopathology in addition to ADHD behaviors, such as anxiety, depression, or atypical behaviors. Examples of such scales are the Child Behavior Checklist (CBCL) and its analogous Teacher Report Form (TRF) (Achenbach, 1991a; Achenbach, 1991b), the Conners' Rating Scales-Revised (CRS-R) (Conners, 1997), and Behavior Assessment Schedule for Children-2nd edition (BASC-2) (Reynolds & Kamphaus, 2004). While the latter rating scales are helpful for gaining general information regarding a wide variety of symptoms related to differential diagnosis, very often when ADHD is suspected, clinicians will use one of the briefer rating scales to assess the corresponding DSM-IV specific symptoms of inattention and hyperactivity/impulsivity.

One such scale used in this study, the ADHD-IV RS (DuPaul et al., 1998), is designed to assess the inattention and hyperactivity/impulsivity symptoms in accordance with the current DSM-IV two-factor dimensional model of ADHD within the home and

school settings. The eighteen symptoms of the DSM-IV diagnostic criteria for ADHD comprise the eighteen items on the scale. Nine inattention items and nine hyperactive/impulsive items are alternated throughout the form (i.e., odd numbers are items inquiring about inattention, even numbers inquire about hyperactivity/impulsivity) so as to reduce rater response bias. The rater is instructed to rate the child's behavior over the past six months (or since the beginning of the school year if the teacher has known the child less than six months) according to a 4-point Likert scale (see Methods for more information). The standardization sample of 2,000 children was selected carefully so as to closely mimic 1990 United States Census information and have similar rates of racial and ethnic group representation. In addition, samples were taken from various regions (e.g., northeast, midwest, south, west) of the United States.

Exploratory factor analyses revealed that a one- or two-factor solution best represented the structure of the scale. Therefore, two subscales were designed to comprise the ADHD-IV RS so as to conform to the two-factor structure of the DSM-IV model of ADHD. Data analyses revealed adequate psychometric properties. Internal consistency values were relatively high (School version: Total score = .94, Inattention = .96, Hyperactivity/Impulsivity = .88; Home version Total score = .92, Inattention = .86, Hyperactivity/Impulsivity = .88), as was test-retest reliability (School version: Total score = .90, Inattention = .89, Hyperactivity/Impulsivity = .88; Home version: Total score = .85, Inattention = .78, Hyperactivity/Impulsivity = .86). Interrater agreement analyses between teachers and parents yielded scores in the moderate range, with coefficient scores as follows: Total score = .41, Inattention = .45, and Hyperactivity/Impulsivity = .40. Further evaluation of the relation between ratings and other criterion measures

revealed that both the Home and School versions of the ADHD-IV RS were significantly correlated with other scales widely used in the assessment of ADHD (e.g., Conners Parent and Teacher Rating Scales). While teacher ratings were also found to correlate significantly with measures of classroom behavior and academic performance, parent ratings were not significantly correlated with these measures, perhaps due to the fact that they are rating the child according to behavior present in a different setting (i.e., home). Lastly, concurrent validity of the scales was found to be adequate as both teacher and parent ratings were demonstrated to differentiate between children with various ADHD subtypes, as well as children in a non-ADHD control group, and the combination of parent and teacher ratings was found to be predictive of ADHD diagnosis (i.e., ADHD group membership) in the context of larger scale clinic and school-based assessments. Therefore, the ADHD-IV RS was found to be a useful diagnostic tool *when used in conjunction* with other measures during the diagnostic process.

Data analyses also revealed that sex differences were found, indicating that boys were rated higher than girls in both the inattentive and hyperactive/impulsive domains. There was also a significant difference between ethnic groups, such that African American children were rated as having more frequent symptoms of ADHD than Caucasian and Latino children. Differences between the Caucasian and Latino groups were variable, with some significant differences in ratings based on older age groups, (e.g., Latino>Caucasian on ADHD ratings in 14-18 year olds) but not in others (i.e., 5-13 years of age). These ethnic group differences held true even after controlling for socioeconomic status. Other ethnic groups were excluded from the analyses due to small

sample size. In addition, due to this small sample, separate norms could not be derived on the basis of ethnic group.

Other rating scales, such as the CBCL (Achenbach, 1991a), TRF (Achenbach, 1991b), the CRS-R (Conners, 1997), and the BASC-2 (Reynolds et al., 2004) differ somewhat in the number of domains they assess (i.e., ADHD-specific behaviors as well as other psychopathology), Likert-scale rating system used (i.e., spread of Likert scale, description of frequency, inclusion of temporal qualifiers), and normative sample. Some scales (e.g., BASC-2) include particularly useful features, such as the inclusion of validity scales to indicate response bias.

All of these scales used relatively large normative samples (between 1,000-5,000 children) but vary in the degree to which they describe the ethnic breakdown of the normative sample. However, in all of the scales, efforts to include children of various ethnicities from various parts of the United States (i.e., east, midwest, south, etc.) were made, with some scales having more varied ethnic representation (BASC-2) than others (CRS-R, CBCL). Analyses of psychometric properties for these scales revealed adequate validity and reliability estimates. Furthermore, all of the scales found both age and gender based differences. However, of concern was the fact that although all of the scales reported at least some differences in ratings based on ethnicity, only some of the scales outlined the nature of the difference in detail (i.e., which ethnic groups differed on which domains within a scale). Scales such as the CSR-R (Conners, 1997) outline specific ethnic group differences found on particular scales, while others such as the CBCL (Achenbach, 1991a) only makes mention of “some scale scores being higher for whites and some for non-whites” (pg. 95).

Taken together, the data from various parent and teacher behavioral rating scales indicate a strong basis for significant age, gender and ethnic group differences. Of particular relevance to this study is the fact that ethnic group differences have, more or less, consistently been noted in the majority of commonly-used rating scales, such that there are significant differences on reported ADHD prevalence/severity ratings between African American and Caucasian groups. In addition, differences between other groups (i.e., Hispanic vs. Caucasian, Asian vs. Hispanic, etc.) are evident in certain ADHD domains depending on the rating scale that is used. However, these findings do not shed light on whether these differences are "real" vs. "reported" (i.e., differences due to some type of bias). Furthermore, since this information is only based on the normative sample studies originally analyzed for use with the corresponding rating scale, it is essential that one look at additional data from research studies which have analyzed differences in prevalence rates of ADHD as a function of ethnicity in more detail.

Differences in prevalence rates across ethnic groups

Issues regarding the validity of various assessment methods are especially critical when considering reported ADHD prevalence rates and whether these results are an accurate reflection (i.e., representative) of the general population. ADHD is not an exception, as the symptomatology may not be similar across different groups and various assessment methods may not be equally valid across groups. Although the DSM-IV-stated prevalence rate of 3-7% is commonly cited throughout the literature, it is oftentimes unclear what type of population was sampled to obtain such a prevalence rate (Rowland, Lesesne, & Abramowitz, 2002). The DSM-IV field trials for ADHD (Frick et

al., 1994; Lahey et al., 1994a; Lahey et al., 1994b) provide few details of sample characteristics related to geographic region, socioeconomic status, or inclusion/exclusion criteria for ethnic groups.

Differences in the prevalence of various psychopathology across cultures have been noted not only within the fields of psychology and psychiatry, but medical anthropology as well (Kleinman, 1988; Marsella & Kameoka, 1989). This point is of particular importance due to the fact that considerable data suggest ADHD prevalence rates vary as a function of ethnic group. Although there is fairly consistent evidence to show that ADHD symptoms exist universally (i.e., cross-culturally), varied prevalence rates have been reported as a function of ethnic group/country. When compared to stated prevalence rates in western societies such as the United States, data suggest higher prevalence rates in German (Baumgaertel et al., 1995), Ukrainian (Gadow et al., 2000), and Columbian (Pineda et al., 1999; Pineda et al., 2003) samples, while data from French (Fombonne, 1994), Icelandic (Magnusson et al., 1999) and Chinese (Leung et al., 1996; Yao, Solanto, & Wender, 1988) populations have reported lower rates of the disorder. Australian (Graetz et al., 2001), Brazilian (Rohde et al., 1999), and Israeli (Margalit, 1981) populations have been found to have similar rates of the disorder when compared to the United States, while data from other countries, such as Italy (Gallucci et al., 1993b; O'Leary et al., 1985a), Puerto Rico (Bauermeister, Berrios, Jimenez, Acevedo, & Gordon, 1990; Bird et al., 1988; Canino et al., 2004) and England (Ford, Goodman, & Meltzer, 2003; Taylor et al., 1984; Rutter, Tizard, Yule, Graham, & Whitmore, 1976), are inconsistent, with lower, similar, and higher reported prevalence rates than that in the United States. See Table 1 for a more extensive review of these data.

Table 1. Review of studies assessing prevalence of ADHD symptoms

Study	Country	Sample Characteristics*	Assessment Methods**	Prevalence of ADHD symptoms ***	General Diagnostic Criteria Model
(Graetz et al., 2001)	Australia	N=3,597 NR, B & G 6 - 17 years of age	DISC-IV parent interview; CBCL; CHQ-PF50 parent version	Without Impairment: 7.5% With Impairment: 6.8%	DSM-IV
(Rohde et al., 1999)	Brazil	N=1,013 NR, B & G 12 – 14 years of age	Screening instrument based on 18 DSM-IV ADHD items; CGAS; CBCL parent and teacher forms; A-PRF	5.8%	DSM-IV
(Leung et al., 1996)	China	N=3,091 NR, B 1st graders	Rutter's Parent (A2) and Teacher (B2) Questionnaires; PACS; CTRS	8.9%	DSM-III-R
(Yao et al., 1988)	China	N=250 NR, B & G 1st – 6th graders	A-TRS; CTRS hyperactivity factor index	10.48% (hyperactive behaviors only)	Mean score of 1.5 or above on Hyperactivity Index or ATRS hyperactivity items
(Pineda et al., 1999)	Columbia	N=540 NR, B & G 4 – 17 years of age	Parent rating scale based on 18 DSM-IV ADHD items	11.3%	DSM-IV
(Pineda et al., 2003)	Columbia	N=330 NR, B & G 4 – 17 years of age	Parent and teacher rating scale screener consisting of 18 DSM-IV ADHD items; Parent and teacher rating scale based on 18 DSM-IV ADHD items; CPRS, CTRS; BASC; Semistructured scale for child psychopathology, parent version	11.5%	DSM-IV
(Rutter et al., 1976)	England	N=2,199 NR, B & G 9 – 11 years of age	Rutter Scales	<1%	?

(Taylor et al., 1984)	England	N=437 NR, B & G 6 – 9 years of age	CTRS; Teacher interview	Mean CTRS hyperactivity scores: Boys: 0.84 / Girls: 0.46	?
(Ford et al., 2003)	England	N=10,438 NR, B & G 5 – 15 years of age	DAWBA parent structured interview; brief DAWBA teacher rating scale	2.23%	DSM-IV
(Fombonne, 1994)	France	N=2,441 NR, B & G 6 – 12 years of age	Parent and teacher survey; CBCL parent version; Rutter (B2) Teacher Questionnaire; Rutter & Graham Parent Interview; CGAS	~1-2%	ICD-9
(Magnusson et al., 1999)	Iceland	N=79 NR, B & G 6 & 8 years of age	ADHD-IV Rating Scale, home and school versions	Teacher rated: 11.7% Parent rated: 9.5%	DSM-IV
(Margalit, 1981)	Israel	N=60 NR, B & G 6 – 12 years of age	Israeli translation of C-ASQ	5% (hyperactive behaviors only)	Score of >15 on C-ASQ
(Gallucci et al., 1993a)	Italy	N=232 NR, B & G 4th graders	Teacher rating scale for screening 14 DSM-III ADHD items	3.9-10.8%	DSM-III-R
(O'Leary, Vivian, & Nisi, 1985b)	Italy	N=344 NR, B & G 2nd – 4th graders	A-TRS	12%	1.5 + Cutoff Score on ATRS
(Baumgaertel et al., 1995)	Germany	N=1,077 NR, B & G 1st – 4th graders	Teacher rating scale consisting of DSM-III-R/DSM IV ADHD and ODD items, as well as some DSM-III-R CD items	DSM-III-R criteria: 9.6% DSM-IV criteria: 17.8%	DSM-III-R DSM-IV
(Bird et al., 1988)	Puerto Rico	N=773 NR, B & G 4 - 16 years of age	CBCL, parent and teacher versions; DISC parent interview; CGAS	Without Impairment: 49.5% With Impairment: 17.9%	DSM-III
(Canino et al., 2004)	Puerto Rico	N=1,886 NR, B & G 4 - 16 years of age	DISC-IV parent interview; PIC-GAS; Service Assessment for Children and Adolescents	Without Impairment: 19.8% With Impairment: 16.4%	DSM-IV

(Gadow et al., 2000)	Ukraine	N=600 NR, B & G 10 – 12 years of age	CSI-4 parent version; CBCL; SCL-90-R; Depression Self-rating Scale for children; Revised Children's Manifest Self-rating Scale; IOWA CTRS	19.8%	DSM-IV
(Nolan, Gadow, & Sprafkin, 2001)	United States	N=3,006 NR, B & G 3 – 18 years of age	ECI-4 teacher version; CSI-4 teacher version; ASI-4 teacher version	15.8%	DSM-IV
(Wolraich, Hannah, Pinnock, Baumgaertel, & Brown, 1996)	United States	N=8,258 NR, B & G K – 5th grade	Modified version of SNAP and Disruptive Behavior Disorder scales	DSM-III-R criteria: 7.3% DSM-IV criteria: 11.4%	DSM-III-R DSM-IV
(Wolraich, Hannah, Baumgaertel, & Feurer, 1998)	United States	N=4,323 NR, B & G K – 5th grade	Modified teacher rating scale of consisting of SNAP, Disruptive Behavior Disorder scales, and Pediatric Behavior Rating Scale	16.1%	DSM-IV

* *Sample Characteristics*: NR= non-referred, R= Referred, B=boys, G=girls.

** *Assessment Methods*: Achenbach- Parent Report Form (PRF); Adolescent Symptom Inventory-4 (ASI-4); Child Behavior Checklist (CBCL); Child Health Questionnaire (CHQ-PF50); Child Symptom Inventory-4 (CSI-4); Conners Parent Rating Scale (CPRS); Conners Teacher Rating Scale (CTRS); Conners- Abbreviated Teacher Rating Scale (A-TRS); Conners- Abbreviated Symptoms Questionnaire (C-ASQ); Development and Well-Being Assessment (DAWBA); Diagnostic Interview Schedule for Children (DISC-IV); Early Childhood Inventory-4 (ECI-4); Inattention/Overactivity With Aggression Conners Teacher Rating Scale (IOWA CTRS); Parent Account of Childhood Symptoms (PACS); PIC-Global Assessment Scale (PIC-GAS); Service Assessment for Children and Adolescents (SACA); Swanson, Nolan and Pelham Rating Scale-Revised (SNAP); Symptom Checklist-90-R (SCL-90-R);

*** *Note*: Study methods and diagnostic criteria vary greatly across studies. When available, key information pertaining to criteria used to obtain prevalence rate (i.e., domain specific behavior, with or without impairment, etc.) is summarized in table. Where prevalence was unavailable, mean scores are cited.

Given that there have been such variable rates reported, epidemiological studies from England and Puerto Rico particularly illustrate the crucial impact of diagnostic criteria, clinician training, and study methodology on reported prevalence rates of ADHD. For example, early studies from England, such as the Isle of Wight studies (Rutter et al., 1976; Rutter, 1989), initially suggested lower rates of ADHD diagnosis than in North American countries. However, those same studies, as well as others to date (Ford et al., 2003; Prendergast et al., 1988; Taylor et al., 1984), have noted that even in the face of lower rates of diagnosis, ADHD symptoms were indeed prevalent in many subjects who were not diagnosed. Furthermore, data suggest that differences in reported prevalence rates between countries may also be due to factors such as variability in diagnostic criteria and clinician training (Prendergast et al., 1988). Findings from Puerto Rican based samples (Bird et al., 1988) have additionally noted the importance of a multi-method approach in assessment of ADHD symptomatology, as studies based on only one measurement (e.g., scores on behavioral ratings) are likely to inflate prevalence rates. In this study, initial prevalence rates of general childhood maladjustment based on behavioral ratings, the Child Behavior Checklist (Achenbach, 1991a), and general DSM-III diagnostic criteria were found to be alarmingly high (49.4%); however, when measures of maladaptive functioning such as the Children's Global Assessment Scale (Bird, Canino, Rubio-Stipec, & Ribera, 1987) were used in conjunction with these methods, prevalence rates dropped from approximately 49% to 18%. This is evident in American samples as well, as studies have found notably higher prevalence rates than the DSM-IV-stated 3-5% when measures of impairment were not taken into account (Nolan et al., 2001; Wolraich et al., 1998).

Taken together, these data suggest that it is difficult to determine the comparability of prevalence rates between countries or ethnic groups, as epidemiological studies oftentimes employ very different samples (i.e., age ranges, socioeconomic status, community vs. clinic-based). For example, considerable data suggest differences in prevalence of psychopathology as a function of the individual being part of a developing vs. developed country (Hopper & Waderling, 2000). Moreover, varied research methods (e.g., rating scale used) and inclusion/exclusion criteria are often used in such studies (Dwivedi & Banhatti, 2005; Faraone, Sergeant, Gillberg C., & Biederman, 2003). In addition, the use of different diagnostic procedures/standards (i.e., DSM-III vs. DSM-IV or DSM-IV vs. ICD) can significantly alter rates of prevalence in the same sample (Gingerich, Turnock, Litfin, & Rosen, 1998; Faraone et al., 2003; Tripp, Luk, Schaughency, & Singh, 1999; Wolraich et al., 1996). Furthermore, epidemiological studies which base prevalence rates on ratings or scores alone, rather than fully incorporating multiple measures which reflect all facets of DSM-IV ADHD (or ICD-10) criteria, are likely to overestimate prevalence of the disorder and further complicate cross-cultural comparisons. In light of the growing evidence of differences in prevalence rates across ethnic groups, and the increasing ethnic heterogeneity of the United States, it is important to not only evaluate the possible confounds in such epidemiological studies, but also to understand the exact reason(s) for discrepancies in prevalence rates across cultures which may be present even when similar study methods are employed. In particular, it is important to know whether *clinical impressions* differ as a function of ethnicity or whether there are *real differences* in the areas of attention, inhibitory control

and activity level in children of different ethnic groups, which may be due to genetic or environmental variation.

The use of ADHD rating scales across ethnic groups

Several studies have attempted to assess the use of behavioral rating scales across ethnic groups. In one of the first reviews which evaluated such cross-cultural differences, Reid (Reid, 1995) introduced the importance of normative and construct (i.e., conceptual) equivalence when using behavioral rating scales with different ethnic groups. Normative equivalence refers to the appropriateness of using the same set of normative data across ethnic groups (i.e., can the same norms be used for various ethnic groups?). If the standard or accuracy by which one measures the occurrence of behaviors is not consistent across ethnic groups, then the scale's normative equivalence would be compromised. Construct equivalence refers to the similarities in conceptual meaning of constructs assessed via behavioral rating scales (i.e., is the scale assessing the same construct across various ethnic/cultural groups?). A common understanding of the actual behaviors and attributes that may make up the symptomatology of a disorder must be present across raters regardless of the ethnicity of the child being rated. If this is not the case, the same construct is not being assessed across groups. In his review of available studies which included data from children of Brazilian, Puerto Rican, American, Chinese, Italian, British, and New Zealand descent, Reid concluded that, generally, the same factor structure was valid for these different ethnic groups, suggesting that construct equivalence was adequate across these ethnic groups. This is consistent with more recent evidence of the same phenomenon. Wolraich et al. (Wolraich et al., 2003) found that the

same two-factor model of ADHD held true across samples from Spain, Germany, urban U.S., and suburban U.S. populations according to teacher ratings on the Vanderbilt ADHD Rating Scale. Similarly, other studies (Beiser, Dion, & Gotowiec, 2000; Yang, Schaller, & Parker, 2000) have found cross-cultural congruency of the two-factor ADHD model when Taiwanese and Native American children were each compared with different U.S. samples on study-designed rating scales which assessed for ADHD symptomatology.

However, as part of his review of studies mentioned above, Reid noted that different cut-off points based on ethnicity should be considered, as some ethnic groups such as African Americans, seem to be over-identified in the literature. Thus, there was not normative equivalence across ethnic groups for most scales evaluated. Of note, the studies reviewed by Reid (Reid, 1995) employed a variety of different rating scales and samples, suggesting that this general finding is consistent across scales. However, it is difficult to compare results of studies with such varied methodological approaches.

To further evaluate whether some ethnic groups seem to be over- versus under-identified, Epstein et al. (Epstein, March, Conners, & Jackson, 1998) studied racial group differences on the Conners Teacher Rating Scale (CTRS) and found that teachers tended to rate African American children higher than Caucasian children on factors relating to externalizing behavior. Furthermore, they found that this was true when rating both boys and girls. Similar to Epstein et al., Reid et al. (Reid et al., 1998) evaluated cross-cultural differences in ADHD as measured by the ADHD-IV Rating Scale (ADHD-IV RS) and found that teachers rated African American students higher than Caucasian students. It should be noted that in this study, the finding held true across all symptoms in both

domains (i.e., inattention and hyperactivity/impulsivity) on the ADHD-IV rating scale, not just one area. This was true across all age groups evaluated (5-18 yrs); however, only a male sample was evaluated, making it impossible to examine possible gender by ethnicity interactions.

Another study evaluating gender and ethnic group differences in behavior as assessed by teacher ratings on the ADHD-IV RS (Reid et al., 2000) supported the general finding of Epstein et al. (Epstein et al., 1998). Factor structure was similar across genders; however, differences were found in ratings across gender and ethnicity such that African American males were rated highest for ADHD symptomatology, followed by African American females and Caucasian males. Caucasian females were rated the lowest. Of significance, externalizing behaviors were found to significantly distinguish Caucasian males and females; however, no such distinguishing difference was found between African American males and females.

Consistent with these findings, Reid et al. (Reid, Casat, Norton, Anastopoulos, & Temple, 2001) found that the two-factor structure of the IOWA Conners Teacher rating scale (i.e., inattention/overactivity, aggression) held true across both European American and African American groups. However, African American boys and girls were rated significantly higher than European American boys and girls, and were 2.5-3.5 times more likely to screen positive for ADHD. Of note in this study, data revealed a rater by ethnicity interaction, such that African American teachers on average rated African American children lower on both scales of the IOWA. In other words, they reported less of a difference between African American and European American children, while European American teachers reported a greater difference between the two groups'

behaviors. Therefore, a European American child would likely be rated lower if rated by a European American teacher than an African American teacher.

The importance of considering such interaction effects (i.e., rater ethnicity by child ethnicity) has been noted in other studies as well. In a sample of children living in the United States, Dominguez de Ramirez and Shapiro (Dominguez de Ramirez et al., 1998) assessed the effects of teacher ethnicity on ratings of Hispanic and "non-Hispanic White" children. Although only non-Hispanic White teachers were selected to rate the children's behavior in this study, their findings emphasized the importance of taking rater ethnicity into account, as their data were contrary to other studies of ADHD symptom prevalence in Hispanic populations, which have noted higher rates of the disorder, but also notably used same-ethnicity raters (i.e., Hispanic). Dominguez de Ramirez & Shapiro used the ADHD-IV RS, the Teacher's Report Form (TRF), and the Conners (Conners, 1989) Abbreviated Teacher Rating Scale (C-TRS) to assess ADHD symptomatology in a large sample of 6-11 year old children. They found significant main effects for gender, such that girls (regardless of ethnicity) were rated as having fewer behavior problems and fewer hyperactive/inattentive symptoms. More interestingly, Hispanics were rated either similar to or lower than non-Hispanic whites on ADHD symptoms across all teacher scales. In no instance were Hispanic children rated as having a higher incidence of ADHD symptomatology when compared to non-Hispanic whites, which is in contrast to previous data (Bauermeister et al., 1990; Pineda et al., 1999; Pineda et al., 2003) which suggest higher teacher-reported ADHD symptoms in children of Hispanic descent. However, the raters in these latter studies were either parents or teachers who were of Hispanic descent as were the children. Thus, it seems

that controlling for rater ethnicity may have an impact on rating, perhaps indicating that the match of rater and child ethnicity changes the perception of the child's behavior. Of note, cultural expectations *within* and *between* cultures may differ, such that one might expect a child of one's own ethnicity to behave in a certain way, but may not hold children of other ethnicities to the same standards, and therefore may bias ratings.

Overall, although the above mentioned studies consistently suggest adequate construct equivalence across ethnic groups, questionable normative equivalence still exists for the behavioral rating scales evaluated (Reid, 1995). However, as noted earlier, true assessment of validity requires not only the administration of the scale of interest, but additionally the administration of an independent criterion measure with which the scale can be compared.

Issues in the assessment of validity

One way of assessing the validity of behavioral ratings is by evaluating whether there is a relationship between the ratings and another test or measure of the same construct. Ideally, this other test or measure should be an accurate and unbiased account of the child's behavior which is independent of the parent and teacher ratings. In other words, one way of evaluating construct validity for behavioral ratings is by looking at the correlation between ratings and an objective measure which assesses that same or a similar behavioral domain.

One of the most unbiased measures of a child's behavior is their performance on laboratory measures, such as computerized tasks which assess, in this case, the constructs of inattention, hyperactivity, and impulsivity. Previous literature has indicated that

children with ADHD exhibit higher rates of errors on computerized laboratory measures such as continuous performance tasks and go/no-go tasks, as well as higher levels of hyperactivity as measured by objective activity monitors when compared to controls (Barkley, 1991; Barkley, 1997; Halperin et al., 1992). Tasks such as these are typically administered via a computer, and presumably provide for a more accurate assessment of hyperactivity/impulsivity and inattention, as there is less subjectivity and/or bias affecting the child's performance.

Nevertheless, the use of objective measures to validate ratings raises important issues regarding the ecological validity of the measures. Specifically, how well do objective measures obtained in laboratory settings represent these same behaviors that are impairing in naturalistic settings? Although some objective measures may lack ecological validity for diagnostic purposes when used alone (Barkley, 1991), others which are correlated with observable behavior can be useful tools which give the clinician more information regarding the validity of other diagnostic measures (Barkley, 1991; Matier-Sharma, Perachio, Newcorn, Sharma, & Halperin, 1995). Certain patterns of poor performance on objective computerized tasks (i.e., CPT tasks & actigraph measures) have been shown to be moderately correlated with ADHD symptomatology (Matier-Sharma et al., 1995). Furthermore, such tasks have been shown to be useful in distinguishing ADHD from controls and some, although not all, other forms of psychopathology (Halperin et al., 1993; Halperin et al., 1992). Thus, data from objective measures can provide valuable information as to whether these "rated" ADHD behaviors are present. Notably, the relatively modest correlations between ratings and objective measures, typically in the .30 - .40 range, are not very different from those reported

between parent and teacher ratings, which also rarely go above .40 (Mitsis et al., 2000; Youngstrom, Loeber, & Stouthamer-Loeber, 2000; Biederman et al., 1993). Yet support for construct validity does not require robust correlations. As long as the correlations are statistically significant, and thus clearly above chance levels, "…even validities as low as .20 or .30 may justify inclusion…" (Anastasi & Urbina, 1997). As such, objective measures can provide a useful way to validate more subjective assessments (i.e., ratings) of related behaviors.

Several possibilities exist in terms of outcome when using objective measures to validate subjective behavioral ratings. For example, a rating scale could be correlated with the criterion measure (and thus valid) for some ethnic groups but not for others, in which case a "slope bias" would exist. Such an outcome is depicted in Figure 1a. In this figure, ratings and objective measures are differentially correlated in ethnicity A and ethnicity B, as indicated by their different slopes. Whereas the data support the validity of the ratings in ethnicity A, the relationship between ratings and measures is minimal in ethnicity B. Therefore the scales should not be assumed to be valid in both groups. Of note, this differential correlation could occur with or without mean differences in ratings or objective measures.

Another possibility is that an "intercept bias" could emerge (see Figure 1b). Here, the two groups differ on ratings, but not on objective measures (or vice versa), but the relationship between ratings and measures is similar in both groups. In this case the scales would be valid for use across ethnic groups, as indicated by similar slopes characterizing the relationship between the objective and subjective measures. However,

due to their differential y-intercept, different cut-off points would be needed such that each ethnic group could be compared to their own appropriate normative group.

Finally, the groups could differ in ratings and objective measures such that the two distributions have the same slope and Y-intercept (i.e., the same regression line) (see Figure 1c). In this scenario, the differences in symptom prevalence as reflected in the behavioral rating scale could be assumed to be due to actual differences in the behavior of the child. Therefore, both ratings and objective measures reflect a real difference in the behavior of children from ethnicity A and ethnicity B.

The use of objective measures as validation criteria

Unfortunately, none of the studies mentioned thus far employed methodology which could allow them to draw conclusions as to whether differences obtained in symptom prevalence on behavioral rating scales reflect real differences in actual child behavior, or instead are a function of the measure (i.e., the rating scale) being differentially valid across ethnic groups. To date, only three studies have attempted to use objective measures when evaluating behavioral rating scales across ethnic groups. Mann et al. (Mann et al., 1992) examined the role of cultural differences in ratings of ADHD-related behaviors in children. In this study, mental health professionals (i.e., psychologists, psychiatrists, social workers) from China, Indonesia, Japan, and the United States rated hyperactive and disruptive behaviors from standardized videotaped vignettes of four boys. Two of the boys were recruited from a clinic while the other two were recruited from a local school. The boys were all 8 years of age, two were from Tokyo and two from Honolulu, and they participated in both individual and group activities.

Each clinician rated each of the four children according to DSM-IIIR criteria for ADHD and the Abbreviated CTRS. The results indicated that reliability of diagnosis within each country was satisfactory; however, a high level of variance was accounted for by the rater's country of residence. Chinese and Indonesian practitioners gave significantly higher scores for hyperactive-impulsive behaviors than did Japanese and American practitioners. However, they did not differ in ratings on the other three videotaped vignettes which contained other disruptive behaviors (i.e., oppositionality, aggression, etc.) Thus, these data provide support for the hypothesis that raters' ethnicity and/or perceptions of children can bias raters' reports of behavior, yet tell us little about the children in regards to symptoms related to hyperactivity.

Sonuga-Barke et al. (Sonuga-Barke, Minocha, Taylor, & Sandberg, 1993) used objective measures to validate teachers' behavioral ratings of English and Asian (i.e., Indian, Pakistani, or Bangladeshi) children. Initially, teachers' ratings of hyperactivity on the Rutter B(2) questionnaire were compared with objective measures of hyperactivity (i.e., actometer readings, number of glances away from a computerized task, number of recorded body movements, and neurologic exam). Results indicated that teacher ratings of Asian and English children did not differ significantly. However, objective measures did not support these "non-differences" in ratings. Instead, Asian "hyperactives" were more similar to English children in the control group on the actometer measures and gazes away from the computerized task. In the second part of this study, they employed a similar procedure; however, they also used classroom observations of the two ethnic groups of children as recorded by one English observer and one Asian observer. Therefore, there was one rater of each ethnic group to match the ethnicity of the children

being rated (i.e., Asian, English) so that one could look at the impact of rater ethnicity on the ratings of the child (i.e., same vs. other ethnicity match). Data about the frequency of gazes off task and body movements were also recorded, and omission and commission errors on a computerized performance task were analyzed. As in the first study, teachers' ratings seemed to overestimate the hyperactive behavior of Asian children. Asian children who were rated as hyperactive were in fact observed to be less active than English children in the classroom (by observers of both ethnicities), which supports the idea that the rater's perception of the child's characteristics, other than the real behavior of the child, can effect rating scale scores. Therefore, even though Asian and English children behaved in similar ways, the level of deviance ascribed to a child seemed to be a function of ethnicity, and therefore Asian children, at least in Britain, may be more likely to be identified as hyperactive. It may be possible that teachers perceive hyperactivity differently in different cultural groups, or rate them according to their expectations of the child's corresponding ethnic group.

A more recent study by Epstein et al. (Epstein et al., 2005) looked at the relationship between teacher ratings and observed classroom behavior in one of the largest ADHD collaborative research samples available, the Multimodal Treatment Study of Children with ADHD. In this study, teachers rated children's behavior on the CTRS, the CBCL, and the SNAP-IV. In addition, children were observed and rated by trained observers according to the Classroom Observation Code (Abikoff & Gittelman, 1985). Of importance, each observer asked the teacher to identify one "comparison child" who was the same ethnicity and gender of the target child, and whose behavior was typical of the other children in the classroom. Only Caucasian and African American participants

were included in the analyses due to small sample size in other ethnic groups. Consistent with previous data, teachers rated African American children significantly higher than Caucasian children. Of note, although classroom observations of children seemed to indicate the same pattern, the higher level of ADHD symptoms in African American children was not significant. Furthermore, the use of classroom observations reduced the difference between the two groups by half. This could be due to several reasons. There may really be a higher prevalence rate of ADHD in African American samples. However, an environmental explanation may also account for the differences, such that African American children may be likely to be in more unstructured and active classrooms, in which case teachers and classroom observations would likely indicate increased severity or number of ADHD behaviors. This could explain why classroom observations reduced the reported symptom rate by half, because the use of comparison children may have indicated that the target child was not in fact deviant from the "norm" of that classroom. Several limitations of this study include the lack of information related to teacher and/or classroom observer ethnicity to assess for rater by child ethnicity interaction effects, as well as the fact that socioeconomic status was not taken into account.

These three studies represent an important first step in the objective assessment of behavioral rating scales across ethnic groups. However, none of the studies systematically determined the nature of the differences as a function of ethnicity, and whether "real" differences truly exist across ethnic groups. In view of the degree to which clinicians rely on subjective ratings from parents and teachers for both diagnostic purposes and to monitor treatment, further research is clearly warranted.

Focus of this study

This study's goal was to use objective measures to validate ADHD ratings. Objective measures of hyperactivity-impulsivity and inattention were used to validate parent and teacher ratings of the child on the ADHD-IV Rating Scale (DuPaul et al., 1998). Specifically, solid-state actigraph data and total errors (i.e., commission plus omission errors) on a CPT/Go-No go task designed for children were used to validate ratings of parent and teacher ADHD symptoms.

With regard to the previous identification of three possible outcomes when using objective measures to validate ratings (see previous section on issues related to validity), it was hypothesized that slope bias would not occur in any of the ethnic groups analyzed. In other words, ratings and objective measures would be correlated in each of the ethnic groups, and therefore adequate construct validity would exist for the scales across ethnic groups. This would provide support for Reid's concept of construct equivalence (Reid, 1995) across ethnic groups. However, it was also hypothesized that evidence for an intercept bias would emerge such that certain ethnic groups would differ from others based on ratings, but not based on objective measures. Specifically, it was predicted that African American children would be rated higher than Caucasian children, and that Caucasian children would be rated higher than Asian children. Hispanic children were expected to be rated similar to Caucasian children. However, no significant differences would be apparent between any of the groups when looking at objective measures. Thus, the ratings would not reflect "true" differences in prevalence rate of these ADHD-related behaviors as they occurred in the laboratory, but rather differences based on perceptions

of the child. This type of finding would ultimately suggest the need for differential cutoffs to be used for children of different ethnic groups. In other words, it was expected that normative equivalence (Reid, 1995) would not be adequate across all ethnic groups.

METHODS

Participants:

Families were recruited as part of a larger study assessing neuropsychological and behavioral predictors of ADHD in preschool children for which approval was granted by the Queens College Institutional Review Board (IRB). The study was fully described to the parents of participating children and one parent for each participating child provided signed informed consent. For that study, three and four year-old children (boys and girls) were recruited from preschools, day care centers, hospitals, health centers/clinics, and doctor's offices within a 10 mile radius of Queens College in Flushing, New York. The larger sample of children consisted of both non-referred (with and without behavioral difficulties) and clinically-referred participants, allowing for wide variation in ADHD symptom levels. However, only non-referred children were used in the present study due to the likelihood of certain ethnic groups being more/less represented in a clinically-referred sample. Exclusion criteria included the child: a) having a chronic medical illness, neurologic disorder, autistic spectrum disorder, b) taking systemic medication, c) having a Full Scale IQ below 80 as assessed by the WPPSI-III, d) not attending school, e) being a non-English speaker. Lastly, since this particular study was focused on validity across ethnic groups, children who did not fit one of the four ethnic group classifications (i.e., African American, European American, Asian, Hispanic) were also excluded.

The final sample for this study consisted of a total of 57 boys, and 29 girls (total N= 86). The sample consisted of 50 European American children (58.1%), 13 Asian children (15.1%), 10 Hispanic children (11.6%), and 13 African American children

(15.1%). A total of 44 children fell into a "Mixed" ethnicity group (i.e., each parent of a different ethnic category), while 8 were of other ethnicities not included in this study (i.e., Middle Eastern, Indian). Furthermore, **7** children who met ethnic inclusion criteria were excluded due to other reasons (e.g., IQ below 80, non-English speaking, neurologic disorder, etc.). Two children were excluded because they were adopted by parents of other ethnicities than the child, in which case parent perception of the child may not be the same as a parent perception of a child of the same ethnicity. Lastly, two children were excluded due to the fact that they were a sibling of another participant in the study in which case the ratings would not be statistically independent of each other. Due to the limited sample size in three of our ethnic groups (i.e., Asian, Hispanic, and African American), our European American group served as a "control" group to which the other three groups were compared. Given the fact that normative groups within the U.S. typically contain a majority of children of Caucasian descent, it would not be unfounded to use our European American group as a control to examine validity of rating scales for other ethnic groups.

Measures:

ADHD-IV Rating Scale (ADHD-IV RS): Parents and teachers completed the Home and School versions, respectively, of the ADHD-IV rating scale (DuPaul et al., 1998). The ADHD-IV rating scale consists of the eighteen DSM-IV symptoms which reflect the two-model factor of ADHD, and asks the parent/teacher to rate the child on each item according to a 4-point Likert scale (0= "never or rarely", 1= "sometimes", 2= "often", 3= "very often").

The Home version of the ADHD-IV RS was standardized on a sample of 2,000 children and adolescents between 4 and 20 years of age. Gender was approximately equally represented in the sample (1,043 girls; 930 boys; 27 unspecified), while analyses for ethnicity revealed that participants were primarily of Caucasian descent (70.2%), followed by African American (15.9%), Latino (5.3%), Asian-American (5.0%), Native American (0.7%), and other/unspecified (3.1%) ethnic origin. The raters on the home version were primarily mothers, with some ratings filled out by fathers, other guardians, or grandparents. Ethnic breakdown for the parents was, as expected, similar to that of the children.

The School version of the ADHD-IV RS was standardized on a sample of 2,000 children and adolescents between the ages of 4 to 19. Gender distribution was relatively even (948 girls; 1,040 boys; 12 unspecified), and prevalence of ethnicity for the children was 65.1% Caucasian, 18.5% African American, 8.0% Latino, 1.7% Asian-American, 0.2% Native American, and other/unspecified 6.6%. As was the case for the child sample, ethnic breakdown of the teachers was similar, with Caucasians (90.2%) largely represented, and African American (6.1%), Latino (1.3%), Asian-American (0.3%), Native American (0.1%), and other/unspecified (2.1%) respectively. Due to restrictions placed by the New York City Department of Education, information regarding the ethnicity of teachers in this study could not be obtained.

Demographics and Development Interview: This parent self-report questionnaire inquired about a variety of demographic characteristics including ethnicity, socioeconomic status, generational status, household composition, and developmental history. Information related to ethnicity and generational status was particularly

important for this study. The ethnic categories listed on the form were as follows: African American, Asian, Caucasian, Hispanic, Native American, Pacific Islander, Mixed Race. This was followed by further inquiry regarding ethnicity (e.g., ascertaining whether participants who classify themselves as Asian are of Japanese, Chinese, Korean, Indian, etc. descent; if Hispanic, whether of Puerto Rican, Dominican, South American, etc. descent). Furthermore, generational status was assessed by inquiry regarding country of birth of the child, of both biological parents, and of both sets of biological grandparents. If any of the above individuals were born outside of the United States, date of move to the United States was obtained. See Appendix A for information relevant to this study collected from the demographics questionnaire.

Solid-state actigraph: Two solid-state actigraphs (Model: WAM-7164) were obtained from Computer Sciences and Applications, Inc. of Shalimar, Florida. The actigraphs were worn around the waist and non-dominant ankle by the children throughout the evaluations to monitor activity level throughout the two sessions. Previous research in school age children found that actigraph data recorded during evaluation sessions are reliable and yield measures that are moderately correlated with parent ($r = 0.29$, $p > 0.05$) and significantly correlated with teacher ($r = 0.32$, $p < 0.01$) ratings of hyperactivity (Reichenbach, Halperin, Sharma, & Newcorn, 1992). Furthermore, in our pilot sample of preschool children ($N = 72$), actigraph counts in our setting were significantly correlated with both parent ($r = 0.28$, $p < 0.05$) and teacher ($r = 0.43$, $p = 0.001$) ratings of hyperactivity/impulsivity on a DSM-IV ADHD checklist.

Kiddie CPT/Go-No go task: This task is a computerized measure (Berwid et al., 2005) developed to assess sustained attention and impulsivity in preschoolers. The task

consisted of two CPT blocks and two Go/No-go blocks to comprise a total of four blocks. Throughout each block, two colorful, engaging stimuli (e.g., flower, bird) were presented on a computer screen with a stimulus duration of 750 milliseconds and an inter-stimulus interval of 1500 milliseconds. One stimulus served as a "go" stimulus (i.e., the child had to press a key) and the other served as a "no-go" stimulus (i.e., the child had to refrain from pressing a key). There were 48 trials (i.e., number of stimuli presented to the child) per each block. The average proportion of go to no-go stimuli varied across four blocks of trials, so that the ratios of go to no-go stimuli were 5:1, 2:1, 1:2, 1:5. The 5:1 and 2:1 ratios constituted a go/no-go measure of inhibitory control and impulsivity. In other words, the majority of the stimuli were targets, with some infrequent non-targets. Therefore, these blocks measured the child's ability to inhibit themselves from responding impulsively and making false alarms. In contrast, the 1:2 and 1:5 ratios constituted a CPT-like measure of sustained attention, as most of the stimuli presented were non-targets. The child therefore had to be vigilant as to when the infrequent target would appear, ideally missing as few as possible. Across all blocks of the tasks, the children had to inhibit their tendency to respond in a particular way (i.e., pressing for non-targets during Go-No go blocks, missing targets during CPT blocks) due to the fact that a prepotent response bias is created in such a task. Each block used separate pairs of stimuli and go and no-go stimuli were counterbalanced across participants. This task took approximately 16 minutes to administer in its entirety. Data from a previous pilot study (Berwid et al., 2005) of these tasks in preschoolers (N = 42) indicated that omission counts were modestly correlated with both parent ($r = 0.29$, $p = 0.058$) and significantly correlated with teacher ($r = 0.34$, $p = 0.019$) symptom counts of inattention. Commission

errors were modestly correlated with both parent (r = .28, p = 0.056) and teacher (r = 0.28, p = 0.056) symptom counts of hyperactivity/impulsivity on a DSM-IV ADHD checklist.

For this task, children were instructed to press a stationary mouse in response to the target stimuli, and to make no response for non-target stimuli. Before each of the four blocks (i.e., 2 CPT, 2 Go-No go), practice trials were administered using a stimulus booklet of 10 stimulus pictures (targets and non-targets) to train the child on the task. This was then followed by a computerized practice trial which consisted of 8 stimuli (target and non-target), during which the children received verbal feedback from the examiner as to whether they made a correct/incorrect response on each trial. Finally, the task was administered for that particular block, during which time the children were not provided with feedback regarding their performance.

Procedures:

Recruitment was completed through a two-stage process. During the first phase of the study, a consent form and parent behavioral rating scale, the ADHD-IV RS (DuPaul et al., 1998), were sent to parents to complete on a voluntary basis through local preschools in the NY area.

Following completion of the parent consent and parent rating scale, the ADHD-IV RS School version was obtained from the child's teacher. Once the teacher scale was obtained, classification into one of two groups was determined. Children who were rated as symptomatic (ratings of 2's or 3's) on at least 6 of the 9 symptoms in either ADHD domain (i.e., 6 of 9 Inattention symptoms or 6 of 9 Hyperactive/Impulsive symptoms) by

either the parent or teacher were assigned to the high risk group. Those children having fewer than 3 ratings of 2's or 3's in both ADHD domains by both raters (i.e., 6 of 9 symptoms in either two domains) were assigned to the low risk group. These criteria based on symptom severity were designed to mirror current DSM-IV criteria for ADHD (i.e., 6 out of 9 symptoms present in either of two domains). Only children who met criteria for either of these groups (i.e., high vs. low risk) were invited into the laboratory at Queens College for further assessment and participation in this study.

Participation in the second phase of the study involved a two day, in-depth psychological evaluation as part of the larger NIH-funded study. Parents who chose to participate in this part of the project were mailed, prior to the first evaluation session, a packet of various questionnaires relating to basic demographic information, as well as parenting style, child behavior, and personality. If forms were not completed by the first evaluation session, parents were asked to complete them while their child was being evaluated. Information regarding ethnic status was obtained through the Demographics and Development Interview. Upon completion of this questionnaire, children were assigned to one of four ethnic groups: African American (not including those participants of Hispanic descent), European American (Caucasian of European descent), Asian (Korean, Chinese and Japanese only), and Hispanic (Dominican, Puerto Rican, Central American, and South American only). It should be noted that the inclusion criteria for each ethnic category yielded groups that were still ethnically heterogeneous. Inclusion and exclusion criteria were created in such a way to minimize heterogeneity, but were made broad enough to have adequate power to detect differences between groups. Both parents had to be of the same ethnicity for the child to be considered that same ethnicity

and be assigned to that ethnic group. In addition, information was obtained regarding generational status (i.e., 1st generation vs. 2nd or 3rd generation) so as to assess degree of acculturation, which might affect the rater's perception of the child. However, generational status did not affect group assignment.

Table 2 shows descriptive statistics for age by total sample, as well as ethnic group. A one-way ANOVA revealed no significant age differences between Asian, Hispanic, and African American groups when compared to the European American group (see Table 2 for statistical significance of each ethnic group compared to European Americans).

Table 2. Age Breakdown by Sample Group

	Mean (years)	**SD**	**Total N**	**T**	**df**	**p**
Total Sample	4.25	0.48	86	--	--	--
European American	4.23	0.50	50	--	--	--
Asian	4.14	0.46	13	0.617	82	>.10
Hispanic	4.34	0.48	10	-0.645	82	>.10
African American	4.37	0.44	13	-0.875	82	>.10

A breakdown of gender and ADHD symptom group status of the total sample, as well as of each ethnic group, can be seen in Figures 2 and 3, respectively. Separate chi-square analyses were used to compare each of three ethnic groups to the European

American group on the basis of gender and group status. Analyses revealed no significant differences in the gender distribution between any of the three ethnic groups when compared to the European American group. However, significant differences (χ^2 = 6.99, df = 1, $p < .01$) were found in the distribution of group status (i.e., number of low risk vs. high risk within each group) between the African American and European American group. No significant differences were found in group status distribution between either the Asian or Hispanic groups in comparison to the European American group.

Figure 2. Gender by Sample Group

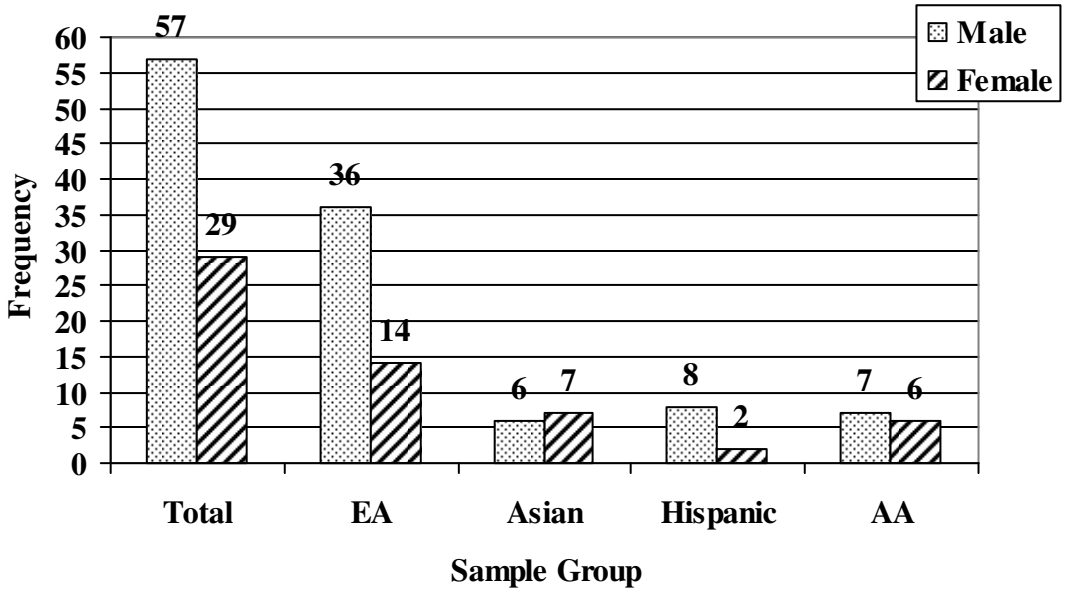
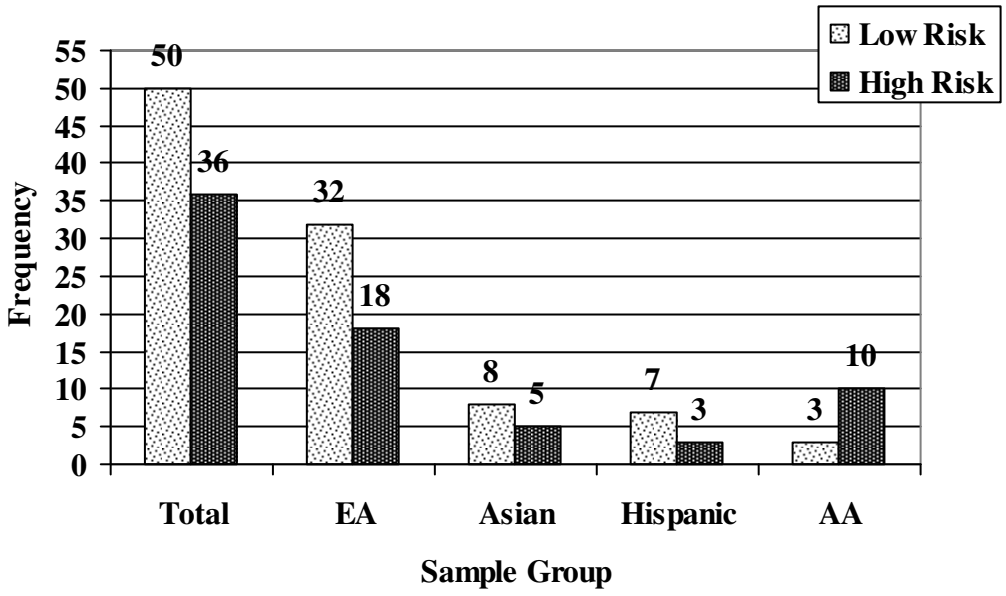


Figure 3. Group Status by Sample Group



Analyses of socioeconomic status (SES) according to the Nakao-Treas Prestige and Socioeconomic Index scores (Nakao & Treas, 1994) yielded SES scores which ranged from 24-89 with a mean of 57.71, indicated most families fell within middle-class socioeconomic status (see Table 3). Results of analyses of variance (ANOVA) revealed the Asian group had a significantly higher SES score than European American group, while the African American group had a significantly lower SES score when compared to the European American group. There was no significant difference in SES between the Hispanic and European American groups.

Table 3. Socioeconomic Status by Sample Group

	Mean (SES score)	SD	Total N	T	df	p
Total Sample	57.71	13.15	86	--	--	--
European American	58.24	13.15	50	--	--	--
Asian	70.77	10.99	13	-2.83	82	<.01
Hispanic	51.30	17.72	10	1.41	82	>.10
African American	47.54	17.80	13	2.42	82	<.05

As part of the larger study, children were evaluated during two, 2 to 3 hour sessions. During these sessions, children were asked to wear two solid-state actigraphs throughout the evaluation to record motor activity. Before the evaluation began, actigraphs were initialized in a reader interface unit, one was placed on the child's waist and non-dominant ankle, and start time (i.e., time of placement) was noted. Any formal breaks (i.e., trip to restroom) were noted and excluded from data analysis. The actigraphs were programmed to record activity level in one-minute epochs for the duration of the child evaluations.

During these sessions, children were administered a lengthy neuropsychological test battery, from which only the Kiddie CPT/Go-No go (Berwid et al., 2005) measures (i.e., total errors) were used for this study. However, other tasks were also administered to the child as part of the battery for the larger study. Testing on the first day included administration of the Wechsler Preschool and Primary Scales of Intelligence-III (WPPSI-III), a video-taped parent/child interaction with structured and unstructured tasks, and the

Kiddie CPT/Go-No go task, respectively. The second day of testing consisted of the NEPSY Developmental Neuropsychological Assessment, interspersed with three computerized tasks inhibitory control, reaction time, and working memory respectively.

At the end of each evaluation session, actigraphs were removed from the child and stop time (i.e., time of actigraph removal) was noted. The actigraphs were then placed in the reader interface unit, stopped, and data were downloaded. After the data were downloaded from the actigraphs to a computer, and imported into a Microsoft Excel spreadsheet, the first five epochs from the start time into the evaluation session, and the last five epochs before the stop time were excluded from analyses so as to minimize bias due to any exaggerated movements from the examiner placing/removing the actigraphs. Mean, median, and standard deviation values were calculated separately for the waist and ankle activity level.

Data Analysis:

Hypothesis 1: *Children would differ as a function of ethnicity on both parent and teacher ratings of behavior as assessed by the ADHD-IV RS. Specifically, African American children would be rated higher than European American children, who in turn would be rated higher than Asian children. Children of Hispanic descent would be rated similarly to European American children. In contrast to parent and teacher ratings of ADHD symptoms, objective CPT total errors, and mean waist actigraph data would not differ as a function of ethnicity.*

Hypothesis 1 was tested using a series of planned comparisons contrasting the European American group with each of the other three groups on parent and teacher ratings as well as objective actigraph and CPT measures. The hypothesis would be supported if the European American group was rated by both parents and teachers significantly lower than the African American group and significantly higher than the Asian group, and if there were no significant group differences on the objective measures. The hypothesis would be further supported by no significant differences between Hispanic and European American children based on parent and teacher ratings, as well as objective measures.

Hypothesis 2: *When objective measures were used to validate parent and teacher ratings, it was hypothesized that slope bias would not exist, and that the scales would be valid across all ethnic groups. Therefore, the subjective and objective measures would be significantly correlated in all groups, and evidence for construct equivalence would be obtained. However, we hypothesized that an intercept bias would emerge in the data. Specifically, it was hypothesized that an intercept bias would emerge for the African American and Asian groups when compared to the European American and Hispanic groups, which would suggest the need for appropriate norms to be determined based on ethnic group for African American and Asian children. Therefore, normative equivalence would not be adequate across ethnic groups.*

The validity of ratings in each group was assessed using one-tailed Pearson product moment correlations to examine the relationship between ratings and objective

measures of behavior. One-tailed tests of significance were used because positive significant correlations would support the validity of the ratings, but negative correlations would be unlikely to occur and would be meaningless other than to suggest the lack of a relationship. However, significant positive correlations in the context of group differences in ratings, but not objective measures, would be indicative of intercept bias and indicate the need for differential norms. Lack of correlation between subjective and objective measures in an ethnic group would indicate lack of validity of the subjective ratings in that ethnic group. Furthermore, if the slope of the regression line from this correlation was different than that of the European American group which we used as a comparison group, this would be indicative of differential validity of the measures (i.e., slope bias) in that ethnic group.

Given the importance of acculturation to the perception of child ethnicity related to this study, information was gathered regarding generational status. In addition, information regarding socioeconomic status was collected. Due to limited power, statistical control for these variables was not possible; however, the vast majority of children in the study were either born in or immigrated to the United States within their first year of life. However, since SES was found to significantly differ between ethnic groups, the data was re-analyzed using SES as a covariate.

RESULTS

Total Sample Characteristics

Table 4 presents descriptive statistics for scores on overall ratings and objective measures for the entire sample, collapsed across ethnic groups. As can been seen, there were adequate ranges for all variables.

Table 4. Descriptive Statistics for Subjective and Objective Measures in Total Sample

	Mean	SD	Range	N
Total Parent ADHD Score	16.11	11.86	0 - 47	84
Total Teacher ADHD Score	14.44	15.16	0 - 52	85
Total CPT errors	38.01	27.13	3 - 119	78
Actigraph Score (waist)	328.32	242.07	13.20 - 1050.96	75

Pearson correlations were calculated to explore the relationships among the four key measures for the total sample (see Table 5). Parent and teacher ADHD ratings were significantly correlated in the total sample, as were total CPT errors with both parent and teacher ADHD ratings. The total number of CPT errors was significantly correlated with mean waist actigraph score. Lastly, mean waist actigraph score was significantly correlated with parent and teacher total ADHD ratings.

Table 5. Pearson Correlations of Subjective and Objective Measures in Total Sample

	Total Parent ADHD Score	Total Teacher ADHD Score	Total CPT errors	Actigraph Score (waist)
Total Parent ADHD Score	1	---	---	---
Total Teacher ADHD Score	.633**	1	---	---
Total CPT errors	.242*	.271**	1	---
Actigraph Score (waist)	.285**	.295**	.233*	1

Note: numbers represent Pearson r value.

** value is significant at the p < .05 level (1-tailed).*
*** value is significant at the p < .01 level (1-tailed).*

European American vs. African American children

Descriptive statistics were calculated for parent total ADHD ratings, teacher total ADHD ratings, CPT total number of errors, and actigraph mean waist score for the European American and African American groups. In general, African American children exhibited higher scores for parent and teacher-rated ADHD symptoms, as well as higher actigraph and CPT error scores. Planned comparisons indicated that the African American group was rated significantly higher on teacher ratings of total ADHD symptoms, with a medium effect size. It should be noted that there was also a trend toward significance (p = 0.056) for differences on parent ratings based on ethnicity, and

results indicated a small effect size. There were no significant group differences on either of the objective measures (see Table 6).

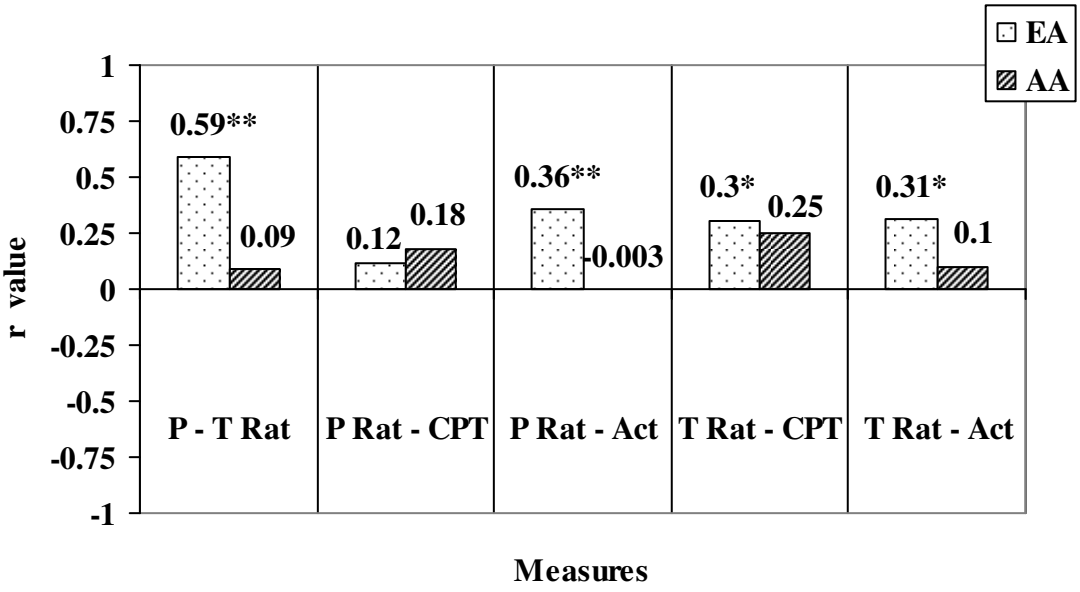
Table 6. Mean Scores of Subjective and Objective Ratings in European American vs. African American Children

	European American		**African American**		**df**	**t**	***p***	**Cohen's *d***
	Mean	SD (Range)	Mean	SD (Range)				
Parent Ratings	14.56	10.88 (0-47)	21.62	10.43 (3-40)	80	1.94	.056	0.43
Teacher Ratings	12.12	13.67 (0-47)	25.77	15.25 (4-52)	81	3.01	.004	0.67
CPT Total Errors	36.88	28.23 (4-119)	45.46	23.36 (14-77)	74	1.00	.319	0.23
Actigraph Waist Score	326.47	249.48 (13.20-1050.96)	332.19	193.65 (47.79-615.00)	71	0.07	.944	0.02

Pearson correlations were calculated to explore the relationship between subjective and objective measures within each of the two groups (see Figure 4). Results revealed significant positive correlations between parent and teacher total ADHD ratings in the European American group. However, there was almost no correlation for the African American group, indicating poor concordance between the two raters (see Figures 4; 5a/5b). Parent ratings were not significantly correlated with CPT total errors

in either of the groups (see Figures 4; 6a/6b). Teacher ratings were significantly correlated with CPT errors in the European American, but not African American group, although the r-values were quite comparable (see Figures 4; 7a/7b) and the discrepancy in significance likely due to issues of power in each group. Similarly, parent ratings and teacher ratings were significantly correlated with actigraph measures in the European American group, but not African American group (see Figures 4; 8a/8b; 9a/9b). Overall, stronger correlations were found between subjective and objective measures in the European American than African American group.

Figure 4. Correlations of Subjective and Objective Measures in European American and African American Children



Note: P Rat = Parent Total ADHD Ratings; T Rat = Teacher Total ADHD Ratings; CPT = CPT Total Errors; Act = Actigraph Score (Waist Mean)

* r value is significant at the $p < .05$ level (1-tailed).
** r value is significant at the $p < .01$ level (1-tailed).

Figure 5a. Parent and Teacher Concordance in European American Children (N=49)

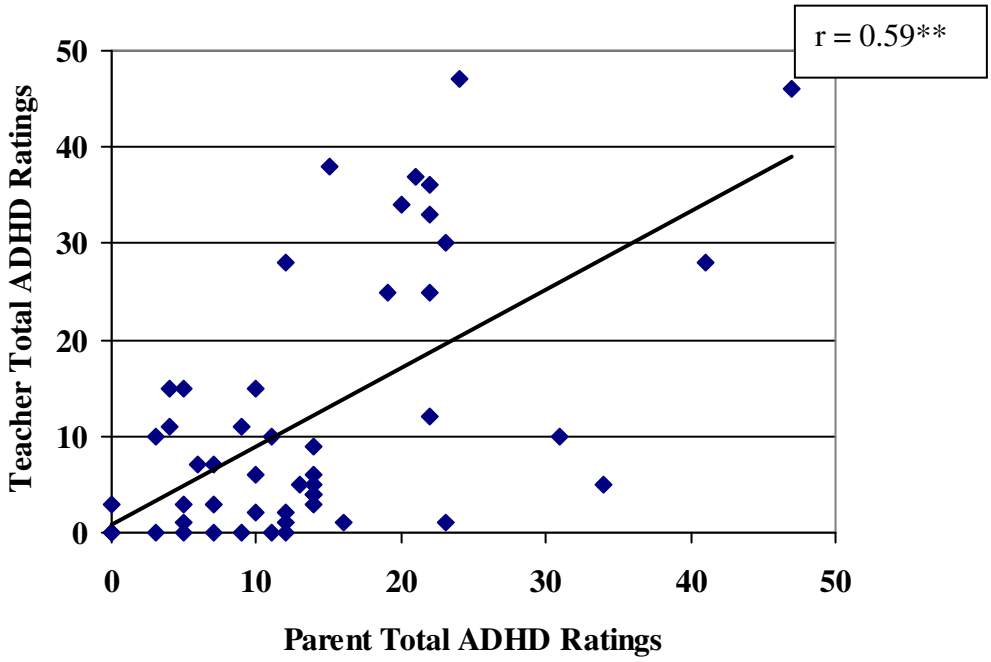


Figure 5b. Parent and Teacher Concordance in African American Children (N=13)

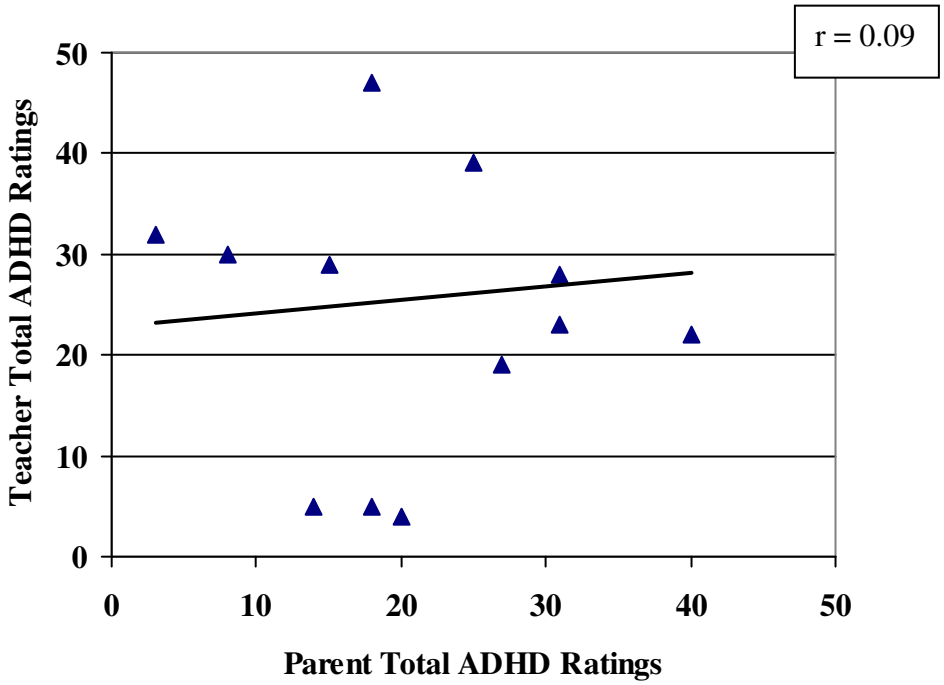


Figure 5c. Parent and Teacher Concordance in Asian Children (N=11)

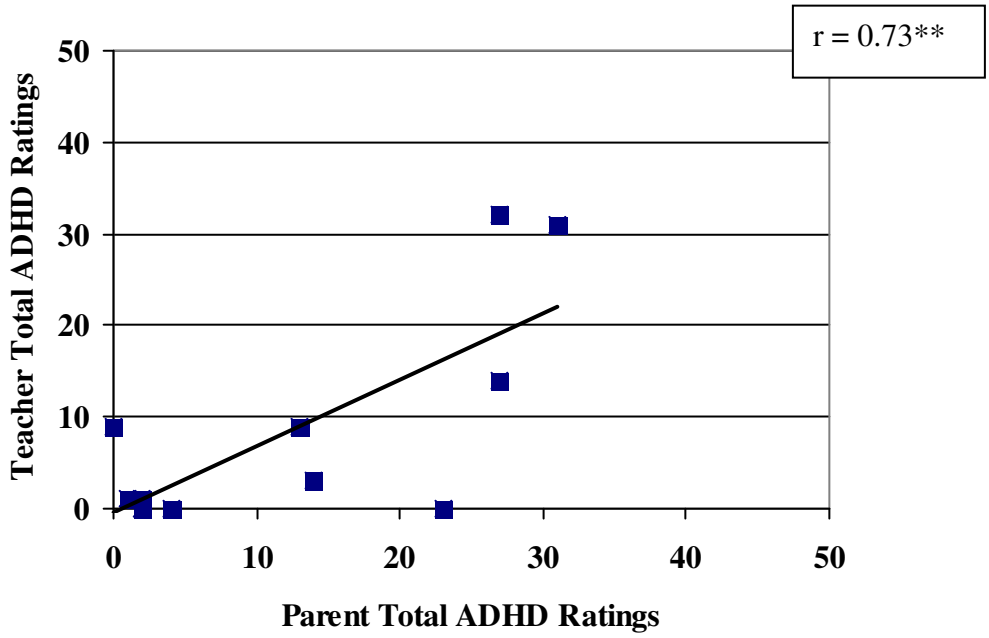


Figure 5d. Parent and Teacher Concordance in Hispanic Children (N=10)

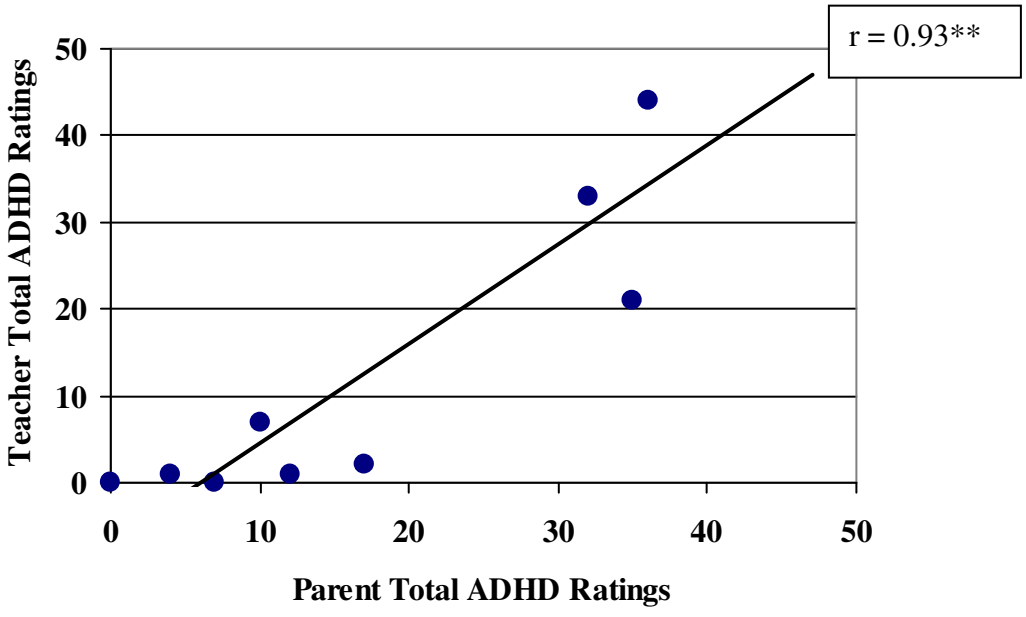


Figure 6a. Parent and CPT Measures in European American Children (N=42)

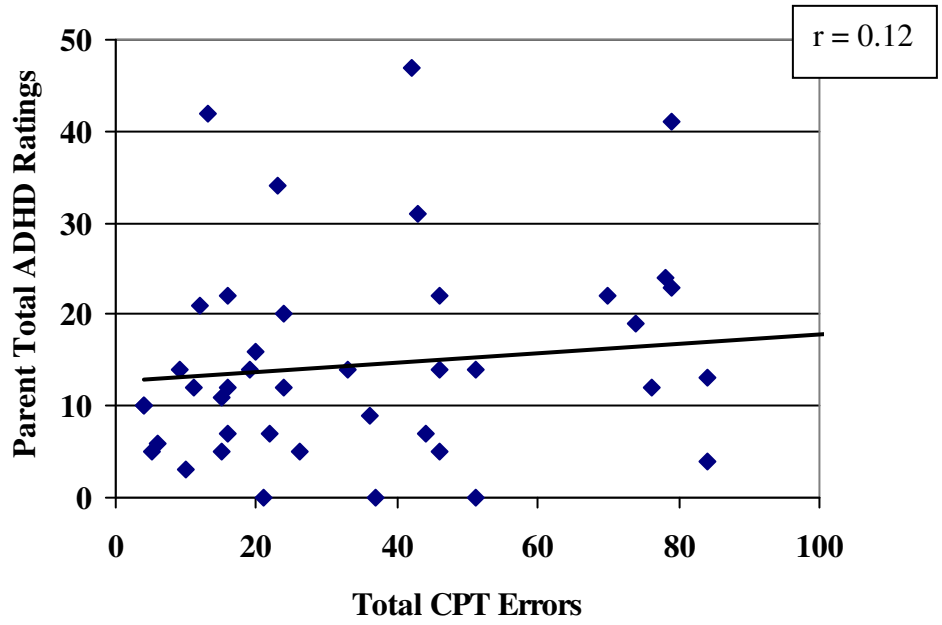


Figure 6b. Parent and CPT Measures in African American Children (N=13)

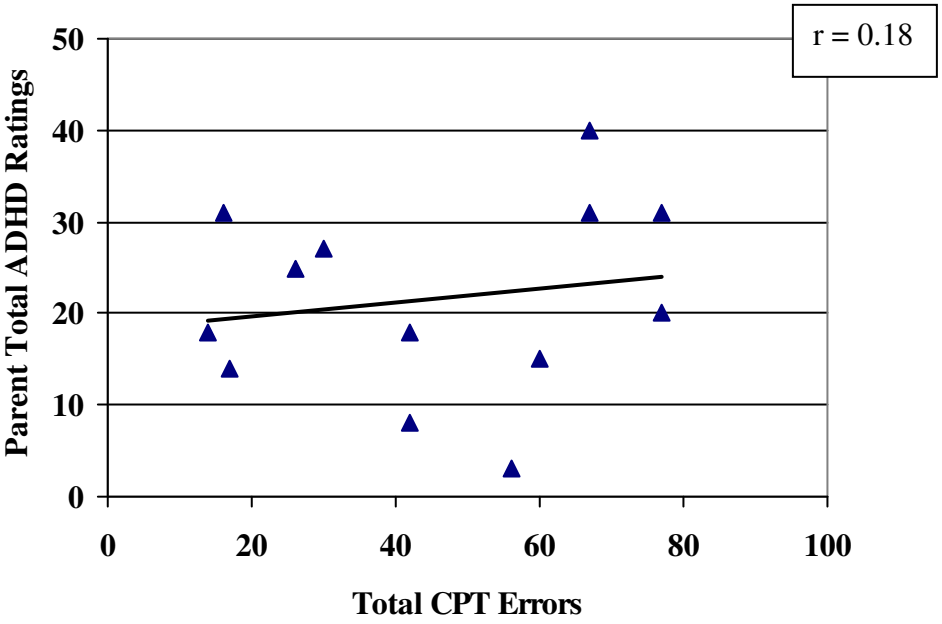


Figure 6c. Parent and CPT Measures in Asian Children (N=11)

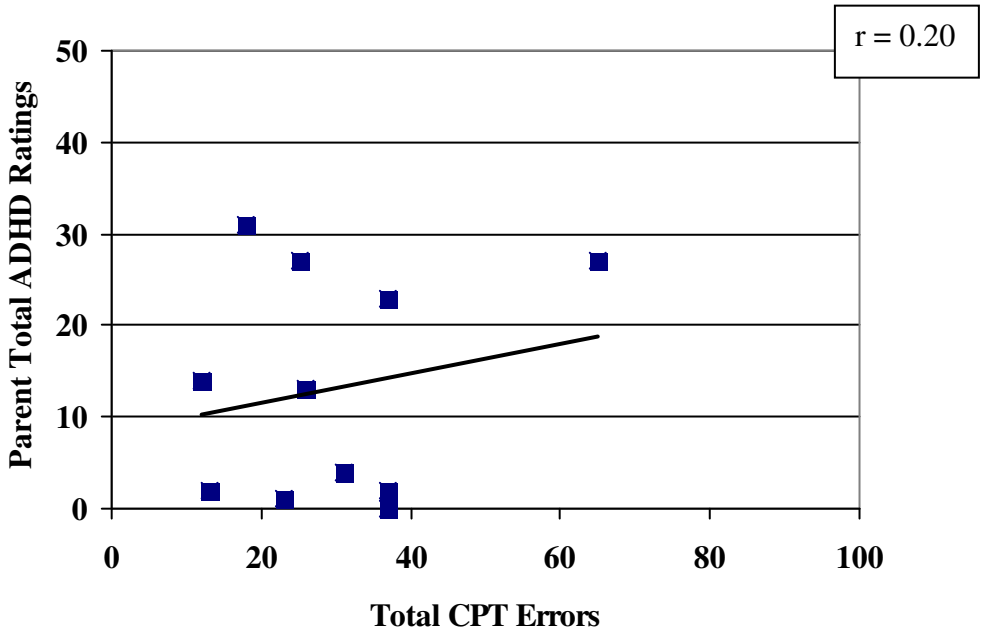


Figure 6d. Parent and CPT Measures in Hispanic Children (N=10)

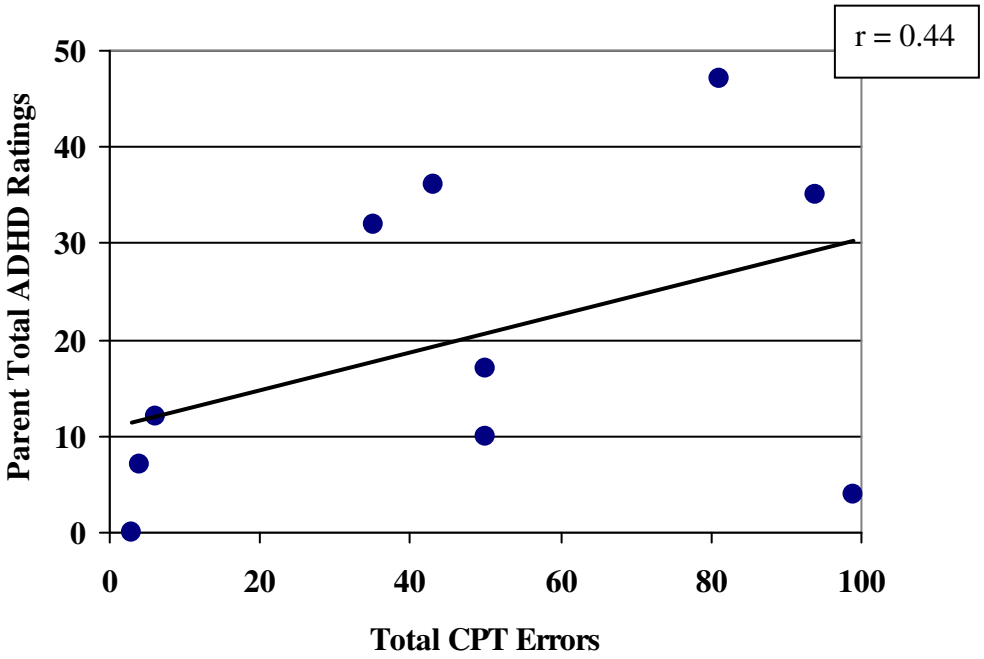


Figure 7a. Teacher and CPT Measures in European American Children (N=42)

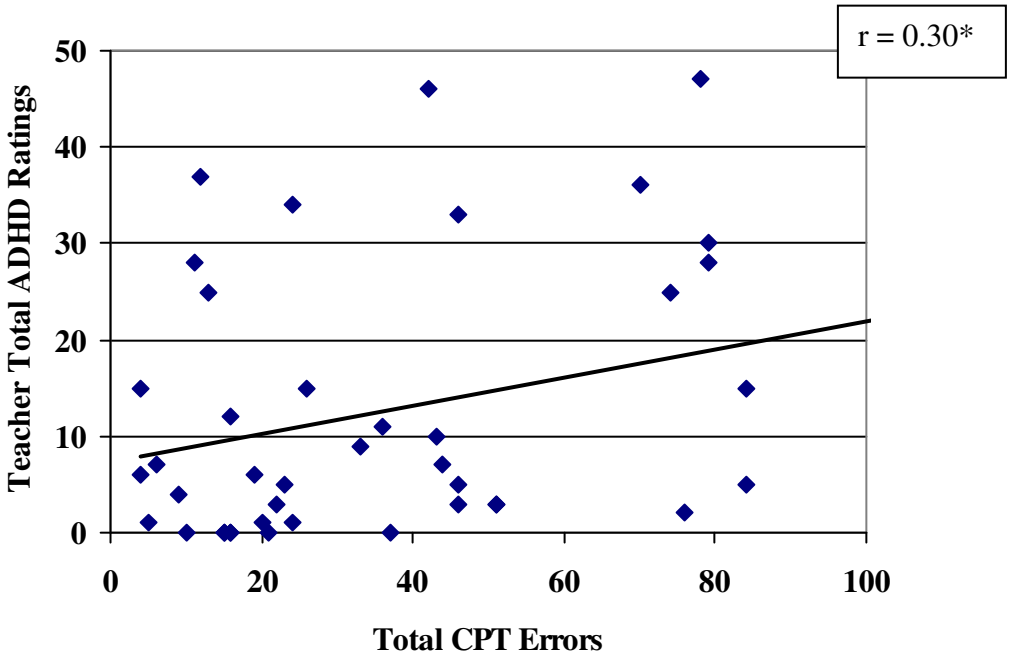


Figure 7b. Teacher and CPT Measures in African American Children (N=13)

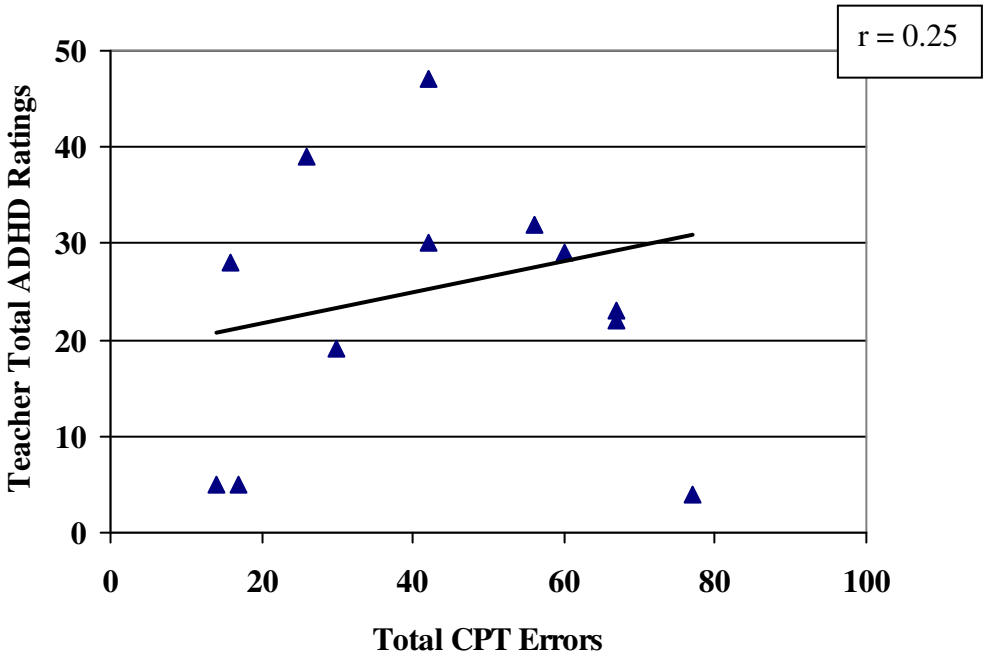


Figure 7c. Teacher and CPT Measures in Asian Children (N=13)

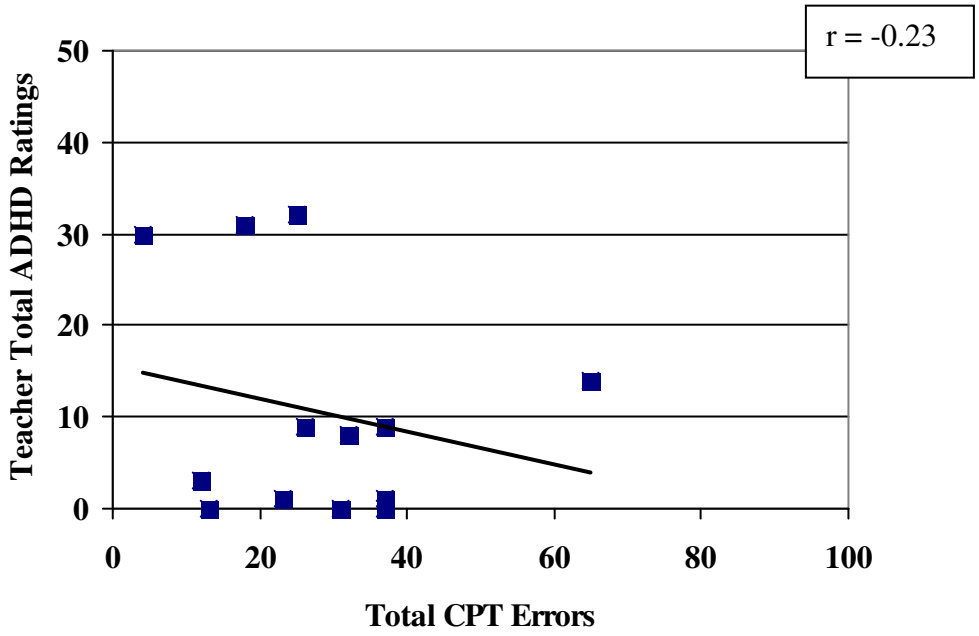


Figure 7d. Teacher and CPT Measures in Hispanic Children (N=10)

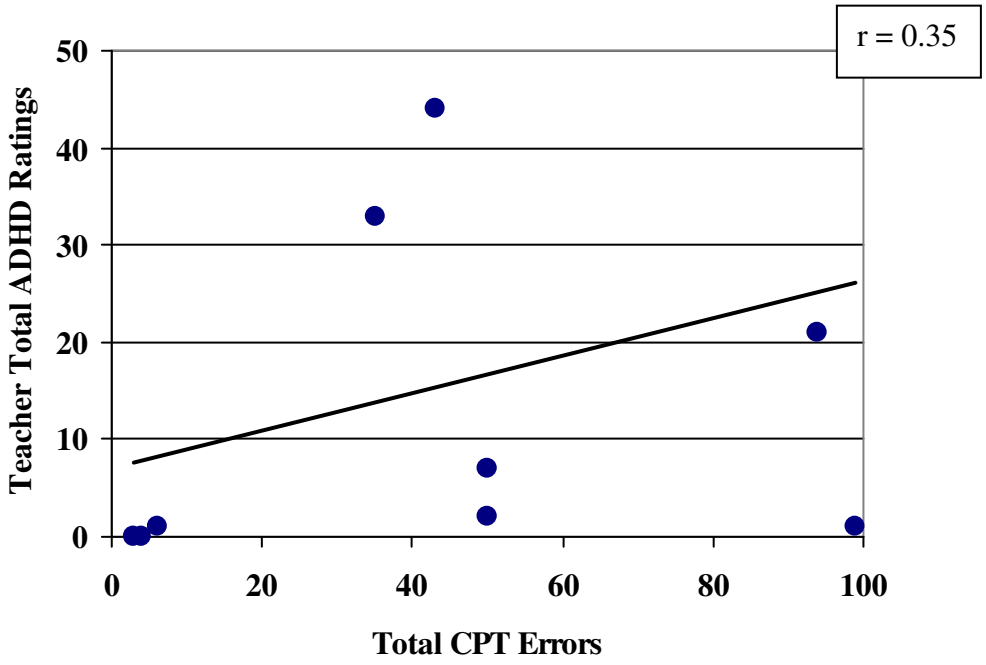


Figure 8a. Parent and Actigraph Measures in European American Children (N=43)

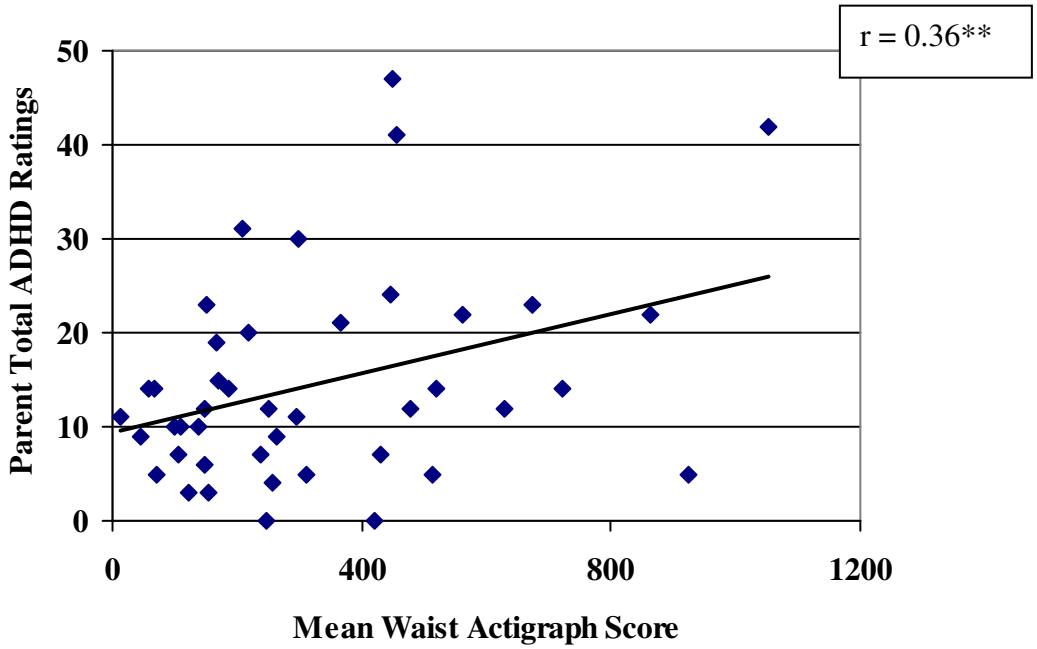


Figure 8b. Parent and Actigraph Measures in African American Children (N=12)

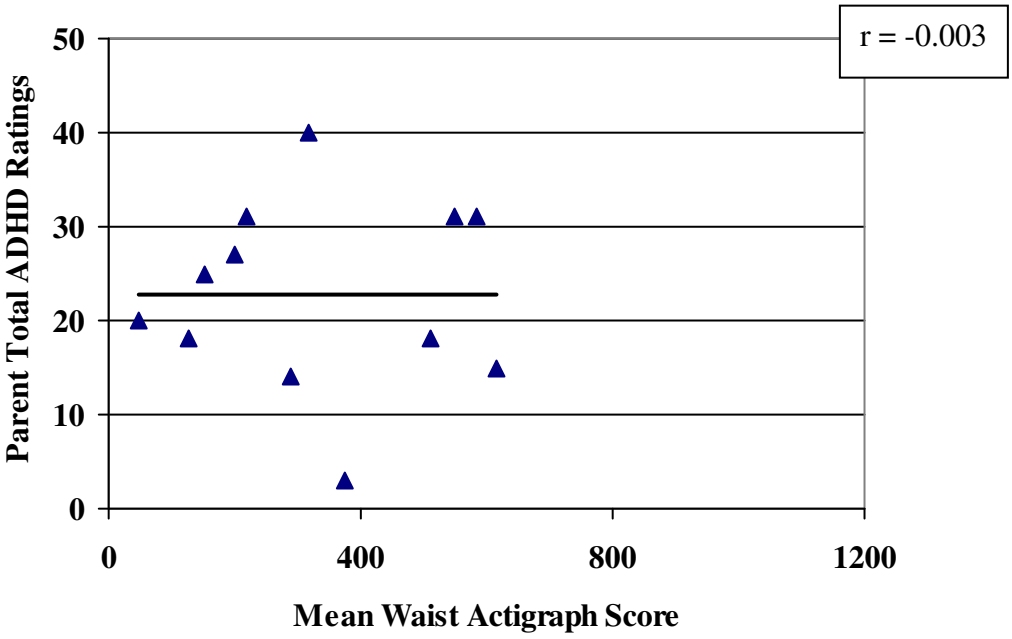


Figure 8c. Parent and Actigraph Measures in Asian Children (N=9)

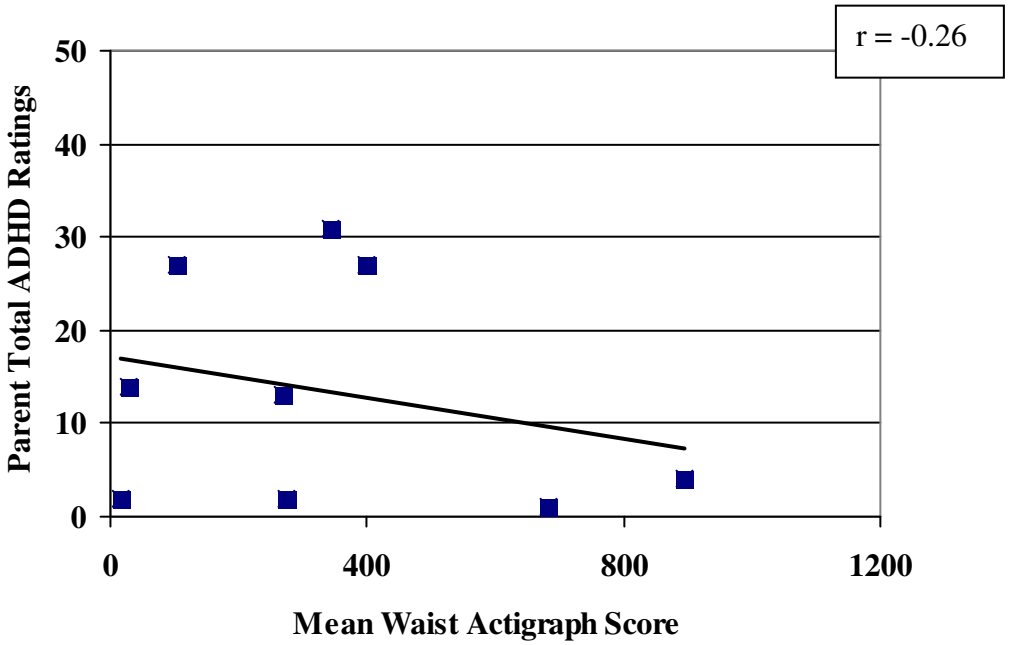


Figure 8d. Parent and Actigraph Measures in Hispanic Children (N=10)

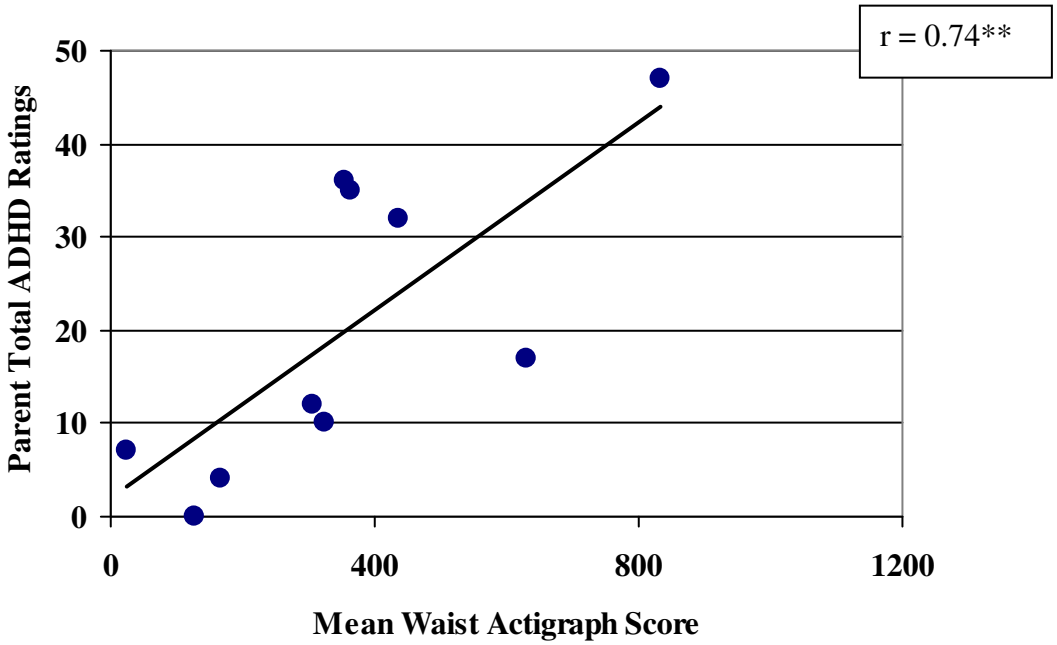


Figure 9a. Teacher and Actigraph Measures in European American Children (N=42)

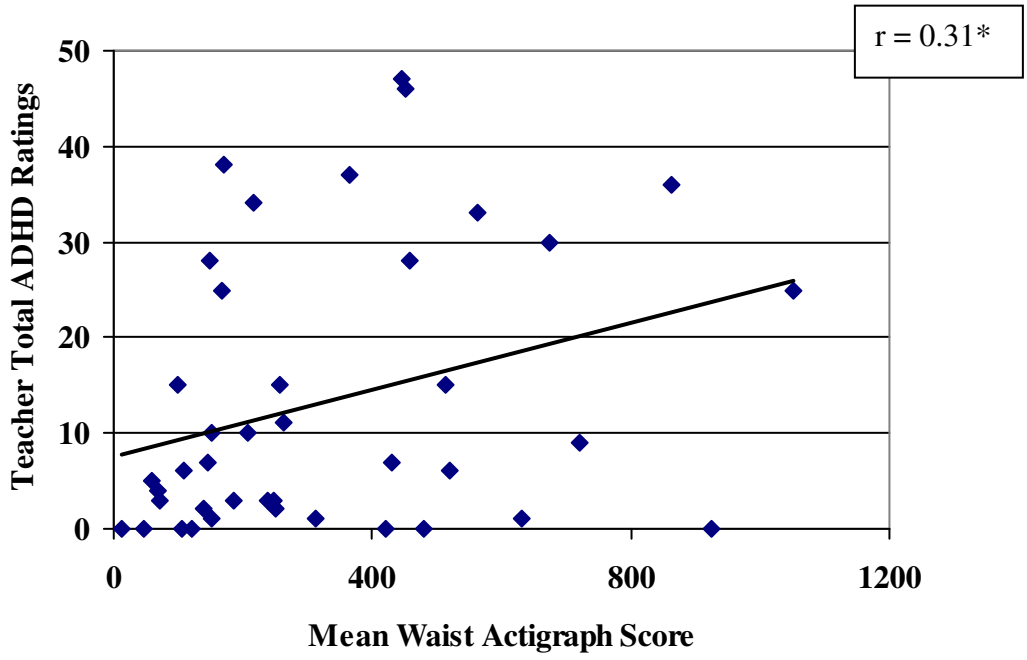


Figure 9b. Teacher and Actigraph Measures in African American Children (N=12)

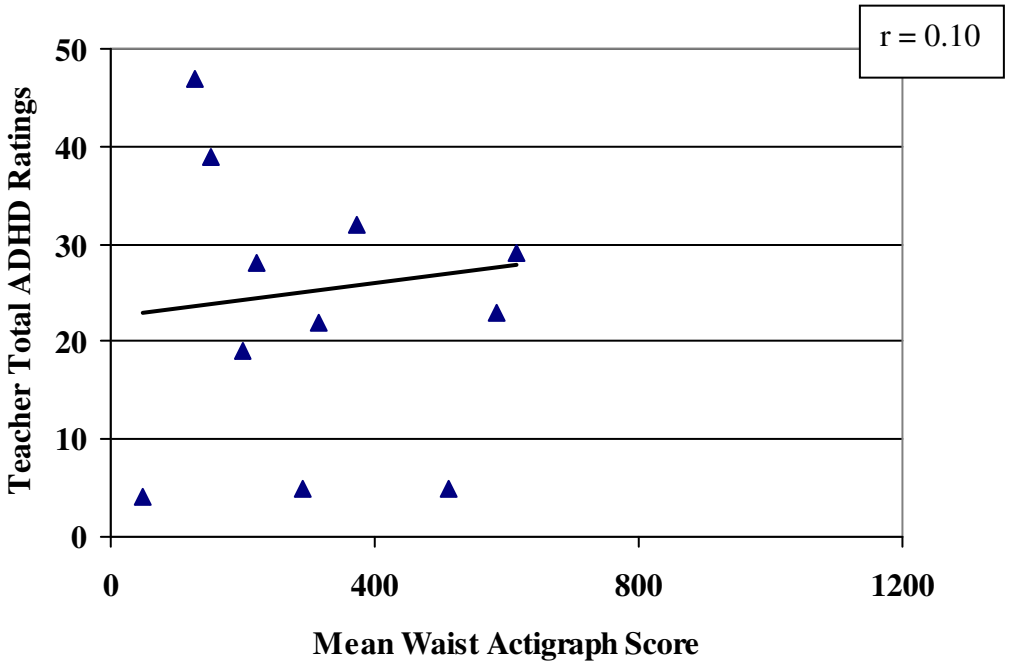


Figure 9c. Teacher and Actigraph Measures in Asian Children (N=10)

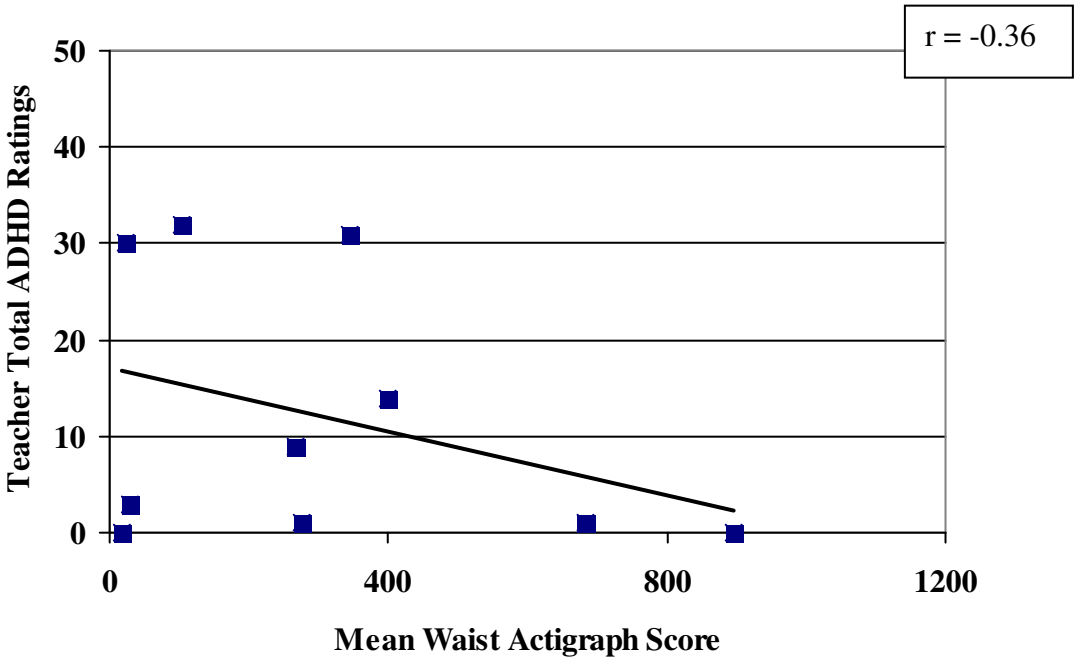
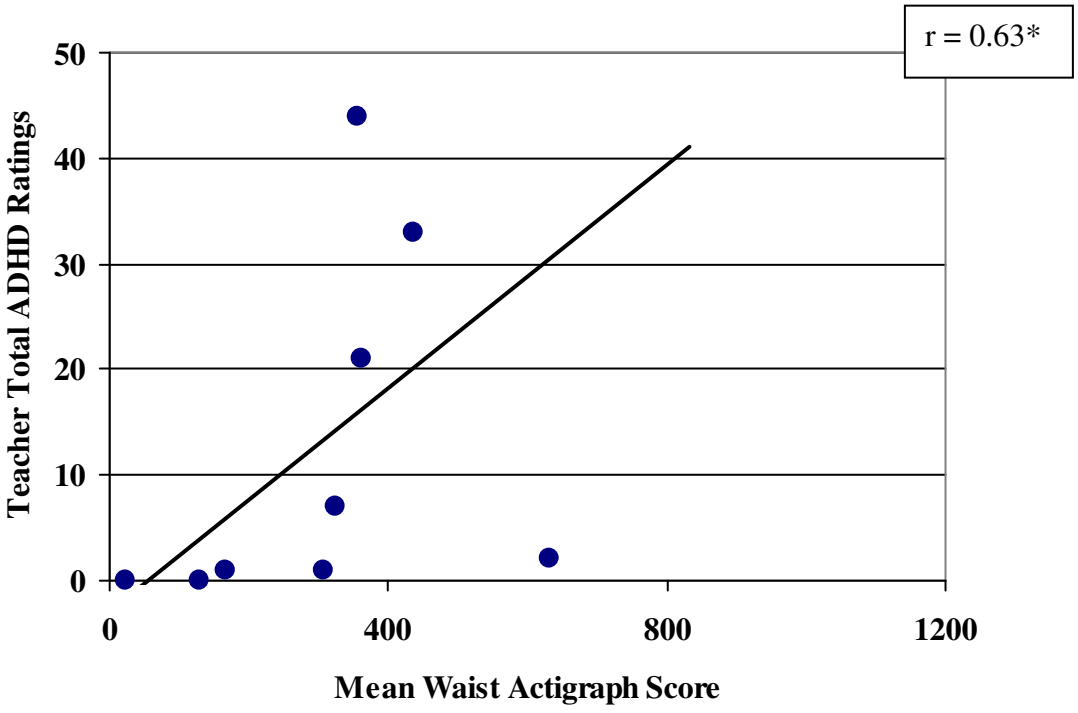


Figure 9d. Teacher and Actigraph Measures in Hispanic Children (N=10)



European American vs. Asian children

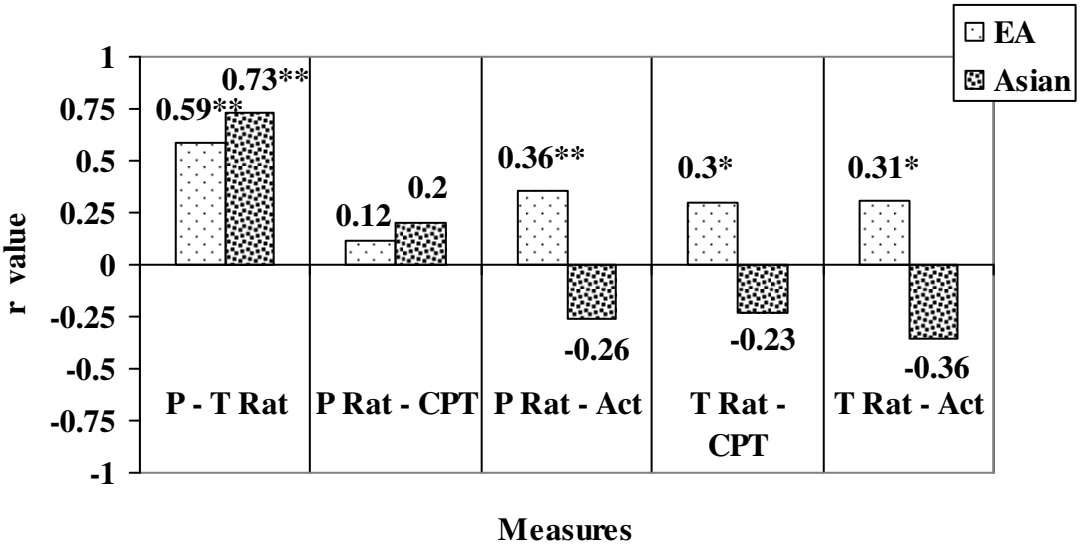
Descriptive statistics for parent total ADHD ratings, teacher total ADHD ratings, CPT total number of errors, and actigraph mean waist score for the European American and Asian groups can be seen in Table 7. Overall, Asians were rated similar to European Americans by both parents and teachers, and had quite similar actigraph scores. Asian children performed somewhat better on the CPT, but this difference was not statistically significant.

Table 7. Mean Scores of Subjective and Objective Ratings in European American vs. Asian Children

	European American		**Asian**		**df**	**t**	***p***	**Cohen's *d***
	Mean	SD (Range)	Mean	SD (Range)				
Parent Ratings	14.56	10.88 (0-47)	13.09	12.04 (0-31)	80	0.38	.707	0.08
Teacher Ratings	12.12	13.67 (0-47)	10.62	12.43 (0-32)	81	0.33	.741	0.07
CPT Total Errors	36.88	28.23 (4-119)	27.69	15.36 (4-65)	74	1.07	.286	0.25
Actigraph Waist Score	326.47	249.48 (13.20-1050.96)	303.44	295.11 (16.79-895.59)	71	0.27	.791	0.06

Pearson correlations indicated significant correlations for parent/teacher concordance in both the European American and Asian groups (see Figures 5a/5c; 10). Specifically, the parents and teachers had particularly good agreement in the Asian group on ratings of ADHD symptoms. Parent ratings were again not significantly correlated with CPT total errors in either of the groups (see Figure 6a/6c; 10). Teacher ratings were significantly correlated with CPT errors in the European American, but not the Asian group (see Figure 7a/7c; 10). Similarly, parent ratings and teacher ratings were significantly correlated with actigraph measures in the European American group, but not Asian group (see Figures 8a/8c; 9a/9c; 10). Of note, a trend toward negative correlations was found on three of the four agreement analyses (parent/actigraph, teacher/actigraph, teacher/CPT) in the Asian group. This would be indicative of slope bias, as both parent and teacher ratings did not accurately reflect objective scores of activity or inattention/impulsivity in Asian children.

Figure 10. Correlations of Subjective and Objective Measures in European American and Asian Children



Note: P Rat = Parent Total ADHD Ratings; T Rat = Teacher Total ADHD Ratings; CPT = CPT Total Errors; Act = Actigraph Score (Waist Mean)

* r value is significant at the $p < .05$ level (1-tailed).
** r value is significant at the $p < .01$ level (1-tailed).

European American vs. Hispanic children

Descriptive statistics for parent total ADHD ratings, teacher total ADHD ratings, CPT total number of errors, and actigraph mean waist score for European American and Hispanic groups can be seen in Table 8. Overall, Hispanic children exhibited somewhat higher actigraph and CPT error scores, and had higher parent and teacher-reported total ADHD symptoms. However, planned comparisons indicated no significant differences between the European American and Hispanic groups on any of the subjective or objective measures.

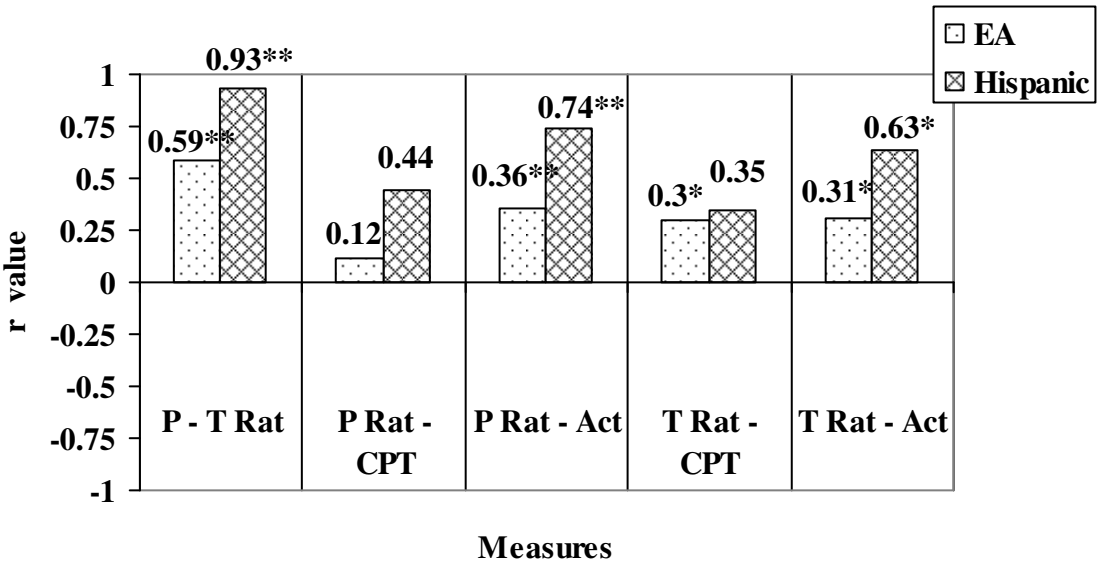
Table 8. Mean Scores of Subjective and Objective Ratings in European American vs. Hispanic Children

	European American		Hispanic		df	t	*p*	Cohen's *d*
	Mean	SD (Range)	Mean	SD (Range)				
Parent Ratings	14.56	10.88 (0-47)	20.00	16.17 (0-47)	80	1.35	.182	-0.30
Teacher Ratings	12.12	13.67 (0-47)	16.00	19.89 (0-51)	81	0.77	.445	-0.17
CPT Total Errors	36.88	28.23 (4-119)	46.50	36.07 (3-99)	74	1.01	.314	-0.23
Actigraph Waist Score	326.47	249.48 (13.20-1050.96)	356.53	238.39 (23-832.54)	71	0.35	.730	-0.08

Pearson correlations indicated significant correlations for parent/teacher concordance for both European American and Hispanic children (see Figures 5a/5d; 11). The Hispanic group had the strongest correlation between parent and teacher ratings of all four ethnic groups. Parent ratings were again not significantly correlated with CPT total errors in either of the groups, although a notably stronger correlation was found between these measures in the Hispanic group (see Figure 6a/6d; 11). Teacher ratings were significantly correlated with CPT errors in the European American group, but not Hispanic group (see Figure 7a/7d; 11). However, the magnitude of the correlations was

quite comparable, indicating similar agreement between the two measures in both groups of children, but a failure to reach significance due to limited number of subjects in the Hispanic group. Parent ratings and teacher ratings were significantly correlated with actigraph measures in both the European American and Hispanic group (see Figures 8a/8d; 9a/9d; 11). Parents, when compared to teachers, had particularly good agreement with objective measures of activity level (i.e., actigraph score) in both groups. In addition, both parents and teachers had relatively stronger agreement with actigraph levels in the Hispanic than in the European American group. These findings were somewhat unexpected in that they were also somewhat indicative of slope bias, such that the Hispanic group had better concordance between subjective and objective measures than the European American group.

Figure 11. Correlations of Subjective and Objective Measures in European American and Hispanic Children



Note: P Rat = Parent Total ADHD Ratings; T Rat = Teacher Total ADHD Ratings; CPT = CPT Total Errors; Act = Actigraph Score (Waist Mean)

* r value is significant at the $p < .05$ level (1-tailed).
** r value is significant at the $p < .01$ level (1-tailed).

Socioeconomic Status

Given the fact that SES significantly differed between some of the ethnic groups, the data were re-analyzed using SES as a covariate. The results of these analyses indicated findings similar to those in which SES was not covaried with one exception. Prior to controlling for SES, only teacher ratings between the European American and African American groups were found to significantly differ. Although parent ratings for these two groups followed a similar trend, the results fell just short of significance (p = .056). After re-analyzing the data using SES as a covariate, these parent ratings were found to significantly differ, such that parents rated their children higher on ADHD symptomatology in the African American group than did parents in the European American group ($p < .05$).

DISCUSSION

This study's aim was to examine the validity of ADHD behavioral ratings across ethnic groups by examining the correlation of subjective parent and teacher ratings to objective measures of theoretically-related constructs. The results of this study were examined in the context of each of three ethnic groups (i.e., Asian, Hispanic, African American) in comparison to a European American ethnic group.

The first hypothesis of this study was that ethnic groups would significantly differ on subjective measures of ADHD symptomatology. Specifically, it was hypothesized that African American children would be rated as having significantly higher symptoms than European American children, Hispanic children would be rated similarly to European American children, and Asian children would be rated as having significantly fewer ADHD symptoms than European American children. Results indicated that African American children were rated significantly higher by teachers, and parent ratings followed a similar trend, falling just short of significance. There was a moderate effect size for differences on teacher ratings between the two groups, while the effect size for parent ratings was small. Thus, both teachers and parents *perceive* African American preschoolers to be more inattentive and hyperactive/impulsive than their European American classmates.

Neither Hispanic nor Asian children were found to significantly differ on either parent or teacher-rated ADHD symptomatology from the European American group. Effect sizes for these measures were generally negligible with the exception of a small

effect size for differences on parent-rated ADHD symptoms between the Hispanic and European American groups.

The second hypothesis of this study posited that no group differences would be found on objective measures of ADHD symptomatology. Results of this study generally supported this hypothesis, as none of the three groups were found to significantly differ on either actigraph scores or CPT errors in comparison to the European American group. However, small effect sizes were found for each of the groups on differences in rates of CPT errors. Specifically, African American and Hispanic children had on average somewhat higher rates of errors than European Americans, while Asians had on average somewhat lower rates of errors than European Americans.

More interestingly perhaps, results of correlational analyses yielded information regarding parent/teacher concordance, as well as each of the raters' concordance with objective measures. While there seemed to be moderate to strong correlations in the European American, Asian, and Hispanic groups respectively for parent/teacher concordance, the agreement between parent and teacher rated ADHD symptoms for African American children was alarming at a value close to zero. This finding is of particular concern in the context of reported higher rates of ADHD symptomatology in African American children (Epstein et al., 1998; Reid et al., 2000; Reid et al., 2001). Given the fact that clinicians oftentimes must rely on both informants to evaluate the cross-situational presence of ADHD, it is disconcerting that data suggest either one or both of the informants (i.e., parents, teachers) are not accurately identifying symptoms in this group of children. This ultimately can lead not only to incorrect prevalence rates, but

also significantly impact the child's quality of life if he/she is not appropriately diagnosed.

On the other hand, it may be the case that these *perceived* parent and teacher differences in child behavior as a function of ethnicity actually exist. In other words, for the African American group, the child's behavior may be perceived differently by the raters as a function of the environment and ethnicity. While both parents and teachers generally rated African American children higher than European American children, within the African American group only, teachers tended to rate children higher than did their parents. The opposite was found to be true in the European American group, as parents generally rated their children higher than did the teachers. Therefore, it could be the case that African American children and European American children differentially exhibit various behaviors across home and school settings which in turn differentially influence parent and teacher perceptions of the child. However, it is important to note that objective measures did not support *real* differences between the two ethnic groups when measured in a laboratory setting. Therefore the differences may be *perceived* differences and not real differences.

Alternatively, it could be the case that there is some inherent bias in the rating scales which is not accurately assessing for the ADHD symptoms in African American children. This is especially likely given the finding of significantly higher teacher ratings, and a trend toward significantly different parent ratings within the context of no significant differences in objective measures of these same ADHD behaviors. However, it is also possible that these children may be exhibiting some behavioral differences that are not necessarily true ADHD symptoms. In other words, African American children

may be exhibiting symptoms similar or related to ADHD (i.e., externalizing behaviors which result in a halo effect and higher ratings on ADHD symptomatology), but are not true ADHD symptoms of overactivity, impulsivity, or inattention that can be assessed by objective measures of these domains. Based on previous literature, this would not be entirely unexpected (Abikoff et al., 1993; Abikoff et al., 2002; Epstein et al., 1998).

A primary goal of this study was to assess the validity of behavioral rating scales by looking at their agreement with objective measures of ADHD behaviors. Results indicated that parent and teacher ratings of European American and Hispanic children positively correlated with objective measures, thus supporting their validity. In contrast, for the African American and Asian groups neither parents nor teachers had adequate agreement with objective measures and therefore the ratings in these groups appear to have questionable validity. In general, actigraph measures yielded higher correlations with parent and teacher ratings than did CPT measures. This may be because participants in this study were preschoolers ages 3-4 years old. In such a sample, it would be relatively more difficult to identify symptoms of inattention, as the majority of children this age present with primarily symptoms of hyperactivity/impulsivity. Our results indicated significant correlations of objective measures of hyperactivity with both parent and teacher ratings which indicate that the overactivity likely present in such an age group is evident across both home and school settings.

Secondly, the Hispanic group, by both parent and teacher, had stronger correlations than any other group for all measures, yielding particularly good validity of subjective measures when compared to objective measures of the same behaviors for these children. Therefore, although the Hispanic and European American groups both

had adequate levels of agreement between most of the subjective and objective measures, the Hispanic group had relatively better agreement between subjective and objective measures than the European American group. This would suggest better validity for the Hispanic children when compared to European American children, and the possibility of slope bias in the European American group.

Of concern was the fact that almost no correlation was found for any of the subjective and objective measures in the African American group, with the exception of questionable agreement between teacher ratings and CPT performance, indicating poor validity of both parent and teacher ratings in this group. Certainly, our finding that African American children were rated higher by teachers, is consistent with previous literature (Epstein et al., 1998; Reid, 1995; Reid et al., 1998; Reid et al., 2000; Reid et al., 2001). However, it is still unclear why there would be such a "reported" difference given the lack of significant findings on objective measures of the same ADHD behaviors. One possible reason for this may be an environmental explanation first proposed by Epstein et al. (Epstein et al., 2005) which posited that African American children are perhaps more likely to be in classrooms in which there is less structure and more activity present due to factors such as higher student/teacher ratios. Epstein et al. (Epstein et al., 2005) found that classroom observations, when used in conjunction with ratings, reduced ADHD symptom prevalence by half. With the integration of classroom observations, the previously significant difference between African American and European American children was found to be non-significant. The classroom observations may reduce bias of ratings because they take into account the fact that the child's behavior is not very

different than the norm, or average child's behavior in that classroom. Therefore, factors inherent in the environment may cause the child's behavior to be perceived differently.

Alternatively, it could be argued that the higher ratings of African American children by teachers were due to the fact that the majority of teachers in our sample were likely to be European American, with a minority of teachers of other ethnic descent. This interaction of rater and child ethnicity is an important issue which has been raised in the literature previously (Dominguez de Ramirez et al., 1998; Reid et al., 2001). However, it is notable that regardless of the discrepancy between raters, parents and teachers both rated African American children higher on ADHD ratings than European American children. Thus, even parents, who are a good ethnic match to their own children, perceived increased rates of ADHD behaviors in African American children. Given the lack of evidence of real differences in objective measures of behavior, this would argue for questionable validity across both parent and teacher ratings in this ethnic group.

Lastly, another interesting finding was the fact that results also indicated a lack of evidence for validity of ratings in the Asian group. In our study, both parents and teachers did not accurately identify ADHD behaviors in Asian children. Given the fact that Asian children qualitatively had the lowest ratings of all four ethnic groups by both raters, but did not differ from the European American group on objective measures, it is likely that some bias is inherent in ratings of behavior for this group as well. It should be noted that this finding may have manifested due to the limited sample size in the Asian group, and may not be truly representative of the whole population. However, if the results are true findings, it provides evidence for slope bias in this group as well,

indicating that both parent and teacher behavioral ratings are invalid for use in the Asian group.

This finding is more difficult to explain in the context of previous studies since, to our knowledge, no such findings address the issue of invalid ratings in Asian children. However, a view into the anthropologic literature may provide more insight into such a phenomenon. Considerable data suggest variability across cultures in the way different groups vary in expression of behaviors (Reid, 1995), perception and/or recognition of normative and non-normative behavior (Kleinman, 1988; Marsella et al., 1989), and treatment options for psychopathology (Halliburton, 2004). One possible explanation for such findings in our Asian group could be that Asian children express core ADHD symptoms in other ways which are evident to the rater, but not evident on objective measures. For example, activity level can vary across cultures, such that children might exhibit increased activity levels during regular play, but not during structured activities. Cultural standards might bias how one perceives that activity, i.e., whether it is normal play vs. overactivity. For example, the importance a cultural group places on obedience and the expectations of performance for the child can play a role in how a child might grow to express inattention. It could be argued that a child adapts to his/her environment such that they learn to express certain behaviors in more subtle ways if they are part of a culture in which obedience, collectivisim (e.g., how the child's behavior reflects on his/her family), and self-control are extremely important (Ho, 1981). These differences between Asian vs. European American cultures have been noted by other researches to possibly affect emergence of hyperactive behaviors in particular (Yao et al., 1988). Others have also noted that, in Chinese groups, inattention seems to be a more prominent

feature of the disorder in comparison to hyperactivity (Tao, 1992). That same finding may hold true for other Asian groups such as those included in this study as well (i.e., Japanese, Korean). It could also be argued that symptoms of inattention are harder to recognize than are symptoms of hyperactivity. If these are the most prevalent symptoms, in the absence of significant hyperactivity, this may lead to invalid ratings as measured by traditional behavioral rating scales.

The concerning lack of validity for parent and teacher ratings in both the African American and Asian groups could also be due to several other factors. One possible explanation is that stereotypes may play a particularly salient role in these ethnic groups and bias ratings in some way. For example, it may be the case that, culturally, parents and teachers have different behavioral expectations or perceptions of children in these ethnic groups. Therefore, the raters may be more likely to misidentify ADHD behaviors in comparison to another ethnic group's normative sample. For example, it may be the case that, given the reported higher prevalence rates of ADHD for African American children (Epstein et al., 1998; Reid et al., 2000; Reid et al., 2001), this ethnic group is likely to be "over-pathologized" and ADHD symptoms over-identified. Similarly, it may be the case that parents and teachers assume Asian children are well-behaved, and not recognize symptoms of ADHD when they indeed are present. However, our data indicate that, at least in the Asian group, both parents and teachers are rating children similarly, which indicates that both raters are susceptible to whatever stereotype may be present. On the other hand, it may that acculturation of Asian children is such that when they come to a new setting (e.g., the laboratory) they become particularly compliant even if that is not characteristic of their typical behavior.

Taken together, this study's findings are not entirely consistent with that of Reid (Reid, 1995). In his review of the literature, Reid concluded that while there was adequate evidence for construct equivalence across ethnic groups, normative equivalence was not acceptable across all groups. Results of this study indicated questionable normative and adequate construct equivalence for two of our ethnic groups (European American and Hispanic). We did not find construct equivalence for African American or Asian groups, as the rating scale used did not seem to be validly assessing ADHD symptomatology in these children (as compared to objective measures). Due to the lack of construct equivalence in these groups, normative equivalence cannot truly be assessed for African American and Asian children. Therefore, our results indicate that for at least some groups, it is not just a basic question of normative equivalence, but a more complex question of whether we are assessing the same ADHD constructs in these ethnic groups.

As noted earlier, much of the literature to date which has indicated quite varied rates of ADHD symptomatology in these (and other) ethnic groups has failed to validate the behavioral ratings with objective measures. Certainly results of this study, as well as the few previous studies (Sonuga-Barke et al., 1993; Epstein et al., 2005) which have employed such validation techniques, strongly indicate the need for further research in this area. As in the present study, future research should aim to create ethnically "meaningful" groups. For example, the ethnic inclusion criteria for this study included a very different sample (i.e., Chinese, Japanese, Korean) than did some Asian groups in previous literature which included those of Indian, Pakistani, and Bangladeshi descent (Sonuga-Barke et al., 1993). That is, inclusion/exclusion criteria should be set such that there is a balance between obtaining an adequate sample size and not overlooking the

differences between ethnic groups even though they may be part of a larger ethnic category (i.e., Indians and Chinese are both Asian, but very different culturally). Data should then be interpreted according to that particular ethnic group, rather than making generalizations to similar, more broad racial categories. Furthermore, there should be differential assessment of subjective and objective measures across these groups so that validity of rating scales can be assessed.

However, there are several limitations to this study which should be taken into consideration. This study included a very small sample size, particularly for making comparisons between four groups. This compromised the statistical power of the study, and made it difficult to interpret certainty of findings. Larger sample sizes would allow for more sophisticated and conservative statistical techniques to be employed, and therefore provide more conclusive results. Due to this limited power, this study also did not control for generational status and degree of acculturation. For example, it could be the case that a family which is more or less acculturated to the "normative" culture perceives and rates their child's behavior differently. This level of acculturation is likely linked to the length of time the family has been in the country, and whether the parents or children are the first, second, third, etc. generation to live in the United States. Information regarding generational status was collected in this study, however the limited power in each ethnic group did not allow for statistical control of this factor. While formal information regarding acculturation was not collected, it is likely that the majority of our families are adequately acculturated, due to the fact that inclusion criteria required at least one parent and the child to be English speaking. Moreover, the vast majority of

children were born in the United States, and attended a preschool, which is likely to suggest not only adequate acculturation of the child, but of the family as well.

Another factor which has been strongly implicated with ethnic differences in psychopathology, including ADHD, is socioeconomic status (SES). Children from low SES families are more likely to come from single parent homes, have more stressful family environments, and be exposed to other environmental factors which are associated with higher prevalence of ADHD (Johnston et al., 2001). Moreover, SES is a variable strongly tied to ethnicity, making certain ethnic groups more susceptible to bias from such a factor. In our analyses, SES was found to differ between Asian and European American groups, as well as African American and European American groups. However, when analyses we re-run using SES as a covariate, the data yielded similar results. Interestingly, only one finding changed as a result of covarying for SES such that the difference between parent ratings in the African American and European American groups now reached significance, whereas previously they just fell short of significance. It is striking that the results of our study generally held true even after controlling for SES. Moreover, controlling for SES yielded a now significant finding. This indicates that, at least in our data set, SES was unlikely to confound results of subjective and objective measures, and that our results reflected differences which could not be accounted for by SES.

While the homogeneity of this sample in regards to generational status, acculturation, and SES is in some ways beneficial statistically for the present study, it may indicate a lack of generalization of findings. In other words, it may be the case that such a study is not sensitive to cultural differences that are present in families of the same

culture who are not as acculturated, whose children are not born here, who are of low SES, or who are non-English speaking. However, it is difficult to study such variables conclusively in cultural research as the subdivisions of cultural, economic, social, and psychological factors which might be relevant to such a study can become virtually endless and impossible to control for.

Furthermore, this study only included preschoolers 3-4 years of age. Given the fact that the majority of ADHD preschoolers are likely to exhibit more symptoms of hyperactivity/impulsivity than inattention, whereas older ADHD children exhibit a more heterogeneous mix of symptoms, it is likely that agreement between objective and subjective measures could be affected by the differential presentation of these behavioral symptoms as a function of age. Generalization of findings to other samples is also limited due to the fact that this study only utilized the ADHD-IV rating scale. Although it is likely that the same results would have been found if another similar scale (i.e., that which assesses the 18 DSM-IV ADHD symptoms specifically) were used, it is possible that another scale could have provided greater sensitivity or specificity in the assessment of ADHD behaviors, therefore affecting the outcome of the subjective ratings and in turn the findings of this study.

Another factor that has been shown to affect accuracy of reported ADHD prevalence in ethnic groups is level of impairment caused by ADHD symptoms (Bird et al., 1988). Therefore, future research in the areas should include a measure of impairment so as not to over-inflate prevalence rates in any particular group. While certain groups of children may exhibit more hyperactive/impulsive or inattentive behaviors, it may not necessarily be impairing. It is possible that the child may not be

rated as highly on subjective behavioral ratings if the behaviors are not perceived as impairing the child's functioning.

Lastly, another limitation of the present study is the fact that there was no information gathered regarding teacher ethnicity. As mentioned previously, rater ethnicity has been shown by this study and others to play an important role in the perception of child behavior (Dominguez de Ramirez et al., 1998; Reid et al., 2001). A study which combines both subjective and objective measures, while controlling for the rater ethnicity x child ethnicity interaction, would provide useful information in delineating the exact role of rater ethnicity in the issue of parent/teacher agreement.

Taken together, our results indicate the universal presence of ADHD symptoms with varied prevalence rates due to differences in research methods and diagnostic criteria (Dwivedi et al., 2005; Faraone et al., 2003). However, data from the present study suggest the possibility of questionable construct and normative equivalence across ethnic groups. Therefore, further assessment of ADHD behaviors in these groups by means of subjective rating scales could provide incorrect information regarding prevalence rates as well as the nature of the disorder in various ethnic groups. Future research in this area is clearly warranted to crystallize the presentation and accurate identification of ADHD symptoms in children of different ethnic backgrounds.

APPENDIX A

PRESCHOOL DEMOGRAPHIC INFORMATION

Child's Name: _____

Date of Birth: ___/___/___ Date of Evaluation: ___/___/___

Age: _____ Sex:_____

Legal Guardian: _____

Physical Guardian: _____

Home Address:_____

Home Phone: _____

Cell Phone: _____

Email: _____

Child's Current School: _____

Grade: _____

Address of School: _____

Respondent:_____

Relationship to Child: _____

Instructions: The items in this questionnaire address issues pertaining to your child's developmental history and family background. For questions that include numbered choice options, please circle the number(s) that best describes your child and/or family. Other items will provide you with space(s) to provide a written response. Be sure to read each item carefully, and direct any questions to a member of the research staff. Try to answer each item as best you can, however, if you feel uncomfortable with any question, you do not need to answer it. Please know that your answers will be kept completely confidential.

I. FAMILY INFORMATION

A. RACIAL BACKGROUND

Please circle *all that apply*:

Background of biological **mother**:

[1] AMERICAN INDIAN / ALASKA NATIVE

[2] ASIAN

[3] NATIVE HAWAIIAN / PACIFIC ISLANDER

[4] BLACK OR AFRICAN AMERICAN

[5] WHITE OR CAUCASIAN

[6] MORE THAN ONE RACE (e.g., White/Black) please indicate:_____

Background of biological **father**:

[1] AMERICAN INDIAN ALASKA NATIVE

[2] ASIAN

[3] NATIVE HAWAIIAN PACIFIC ISLANDER

[4] BLACK OR AFRICAN AMERICAN

[5] WHITE OR CAUCASIAN

[6] MORE THAN ONE RACE (e.g., White/Black) please indicate: _____

B. ETHNIC BACKGROUND

Please indicate with a check mark *all that apply*:

Background of biological **mother**:

Hispanic or Latino?
[YES] [NO]

If ***Hispanic or Latino***,
indicate whether of:

____ Puerto Rican descent
____ Dominican descent
____ South American descent
____ Other, please indicate:

Background of biological **father**:

Hispanic or Latino?
[YES] [NO]

If ***Hispanic or Latino,***
indicate whether of:

____ Puerto Rican descent
____ Dominican descent
____ South American descent
____ Other, please indicate:

If not Hispanic, please check all that apply below:

If ***Asian***, indicate whether of*:*

____ Korean descent
____ Chinese descent
____ Japanese descent
____ Indian descent
____ Other, please indicate:

If ***Asian***, indicate whether of:

____ Korean descent
____ Chinese descent
____ Japanese descent
____ Indian descent
____ Other, please indicate:

If ***Black or African American****,* indicate whether of:

____ African descent
____ Caribbean descent
____ Other, please indicate:

If ***Black or African American***, indicate whether of:

____ African descent
____ Caribbean descent
____ Other, please indicate:

If ***White or Caucasian***, indicate whether of:

____ European descent
____ Middle Eastern descent
____ Other, please indicate:

If ***White or Caucasian****,* indicate whether of:

____ European descent
____ Middle Eastern descent
____ Other, please indicate:

For the following family members, please list their country of birth. If the individual was not born in the U.S., please indicate the year they moved to the U.S.:

FAMILY MEMBER	**COUNTRY OF BIRTH**	**YEAR OF MOVE** ***(if not applicable, please write N/A in the space below)***
Child		
Mother		
Father		
Maternal Grandmother		
Maternal Grandfather		
Paternal Grandmother		
Paternal Grandfather		

II. HOUSEHOLD INFORMATION

Marital Status of biological parents *(circle one)*:

[1] MARRIED
[2] MARRIED BUT SEPARATED; date of separation _____
[3] DIVORCED; date of divorce _____
Mother: single____ remarried____; date of remarriage _____
Father: single____ remarried____; date of remarriage _____
[4] NEVER MARRIED/SINGLE

If SEPARATED OR DIVORCED, specify:

Mother's contact with child:
[1] DAILY
[2] WEEKLY
[3] MONTHLY
[4] YEARLY
[5] NONE

Father's contact with child:
[1] DAILY
[2] WEEKLY
[3] MONTHLY
[4] YEARLY
[5] NONE

Is child adopted or a foster child? *(circle one)* [1] NO [2] YES
If yes, age at adoption or placement: _____

Please list all household members:

FULL NAME	SEX	AGE	RELATION TO CHILD If not FULL siblings indicate relatedness: FULL, STEP, HALF, ADOPTIVE	OCCUPATION

III. INCOME/EDUCATION INFORMATION

Major source of Household Income:

[1] PATERNAL EMPLOYMENT

[2] MATERNAL EMPLOYMENT

[3] COMBINED PARENTAL EMPLOYMENT

[4] OTHER'S EMPLOYMENT

specify whose: _____

[5] ENTITLEMENTS *(Pensions, S.S.I., etc.)*

specify: _____

[6] PUBLIC ASSISTANCE

[7] OTHER

specify: _____

Is **father** currently employed? *(circle one)* [1] NO [2] YES [99] DON'T KNOW

If YES,

Current employer: _____

Date started: _____

Current position: _____

Is **mother** currently employed? *(circle one)* [1] NO [2] YES [99] DON'T KNOW

If YES,

Current employer: _____

Date started: _____

Current position: _____

Reference List

1. Abikoff, H., Courtney, M., Pelham, W. E., & Koplewicz, H. S. (1993). Teachers' ratings of disruptive behavior: Influence of halo effects. *Journal of Abnormal Child Psychology, 21,* 519-534.

2. Abikoff, H. & Gittelman, R. (1985). Classroom Observation Code: A modification of the Stony Brook code. Psychopharmacol.Bull. 21, 901-907.

3. Abikoff, H. B., Jensen, P. S., Arnold, L. L., Hoza, B., Hechtman, L., Pollack, S. et al. (2002). Observed classroom behavior of children with ADHD: relationship to gender and comorbidity. *Journal of Abnormal Child Psychology, 30,* 349-359.

4. Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist 4/18 and 1991 Profile*. Burlington: University of Vermont.

5. Achenbach, T. M. (1991b). *Manual for the teacher's report form and 1991 profile*. Burlington: University of Vermont.

6. AmericanPsychiatricAssoc. (1994). *Diagnostic and statistical manual of mental disorders (4th ed)*. (4th ed.) Washington,D.C.: American Psychiatric Association.

7. Anastasi, A. & Urbina, S. (1997). *Psychological Testing*. (7 ed.) Upper Saddle River, New Jersey: Prentice Hall.

8. Antrop, I., Roeyers, H., Oosterlaan, J., & Van Oost, P. (2002). Agreement between parent and teacher ratings of Disruptive Behavior Disorders in children with clinically diagnosed ADHD. *Journal of Psychopathology and Behavioral Assessment, 24,* 67-73.

9. Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology, 19,* 149-178.

10. Barkley, R. A. (1997). *ADHD and the nature of self-control*. New York: The Guildford Press.

11. Barkley, R. A. (2000). *Taking Charge of ADHD: The complete authoritative guide for parents.* (Revised ed.) New York: The Guilford Press.

12. Barkley, R. A., Cook, E., & et al. (2002). International Consensus Statement on ADHD. *Clin.Child Fam.Psychol.Rev., 5,* 89-111.

13. Bauermeister, J., Berrios, V., Jimenez, A., Acevedo, L., & Gordon, M. (1990). Some Issues and Instruments for the Assessment of Attention-Deficit Hyperactivity Disorder in Puerto Rican Children. *Journal of Clinical Child Psychology, 19,* 9-16.

14. Baumgaertel, A., Wolraich, M. L., & Dietrich, M. (1995). Comparison of diagnostic criteria for attention deficit disorders in a German elementary school sample. *J.Am.Acad.Child Adolesc.Psychiatry, 34,* 629-638.

15. Beiser, M., Dion, R., & Gotowiec, A. (2000). The structure of attention-deficit and hyperactivity symptoms among Native and non-Native elementary school children. *Journal of Abnormal Child Psychology, 28,* 425-437.

16. Berwid, O., Curko Kera, E. A., Marks, D., Santra, A., Bender, H., & Halperin, J. (2005). Sustained attention and response inhibition in young children at risk for Attention Deficit/Hyperactivity Disorder. *Journal of Child Psychology and Psychiatry, 46,* 1219-1229.

17. Biederman, J., Faraone, S. V., Doyle, A., Lehman, B. K., Kraus, I., Perrin, J. et al. (1993). Convergence of the Child Behavior Checklist with structured interview-based psychiatric diagnoses of ADHD children with and without comorbidity. *J Child Psychol Psychiatry, 34,* 1241-1251.

18. Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention-deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *American Journal of Psychiatry, 148,* 564-577.

19. Bird, H., Canino, G., Rubio-Stipec, M., Gould, M., Ribera, J., Sesman, M. et al. (1988). Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico: The use of comorbid measures. *Arch Gen Psychiatry, 45,* 1120-1126.

20. Bird, H., Canino, G., Rubio-Stipec, M., & Ribera, J. (1987). Further measures of the psychometric properties of the Children's Global Assessment Scale. *Arch Gen Psychiatry, 44,* 821-824.

21. Campbell, S. B. (2002). *Behavior problems in preschool children: Clinical and developmental issues*. New York: The Guilford Press.

22. Canino, G., Shrout, P. E., Rubio-Stipec, M., Bird, H. R., Bravo, M., Ramirez, R. et al. (2004). The DSM-IV rates of child and adolescent disorders in Puerto Rico: prevalence, correlates, service use, and the effects of impairment. *Arch.Gen.Psychiatry, 61,* 85-93.

23. Conners, K. (1989). Conners' Teacher Rating Scales manual. North Tonawanda, NY, Multi-Health Systems.

24. Conners, K. (1997). Conners Rating Scale Revised: Technical Manual. New York, Multi-Health Systems.

25. Dominguez de Ramirez, R. (2001). *Effects of student ethnicity on judgements of ADHD symptoms among Hispanic and non-Hispanic white teachers.* PhD Lehigh University.

26. Dominguez de Ramirez, R. & Shapiro, S. (1998). Teacher ratings of Attention Deficit Hyperactivity Disorder symptoms in Hispanic children. *Journal of Psychopathology and Behavioral Assessment, 20,* 275-293.

27. DuPaul, G., Guevremont, D. C., & Barkley, R. A. (1991). Attention Deficit-Hyperactivity Disorder in Adolescence: Critical Assessment Parameters. *Clinical Psychology Review, 11,* 231-245.

28. DuPaul, G., Power, T., Anastopoulos, A., & Reid, R. (1998). ADHD Rating Scale--IV: Checklists, norms, and clinical interpretation. 0-79. New York,NY, Guilford Press.

29. Dwivedi, K. N. & Banhatti, R. G. (2005). Attention deficit/hyperactivity disorder and ethnicity. *Archives of Dis.Child, 90,* i10-i12.

30. Epstein, J. N., March, J. S., Conners, C. K., & Jackson, D. L. (1998). Racial differences on the Conners Teacher Rating Scale. *Journal of Abnormal Child Psychology, 26,* 109-118.

31. Epstein, J. N., Valencia, E. Y., Tonev, S. T., Arnold, L. E., Willoughby, M., Abikoff, H. B. et al. (2005). The role of children's ethnicity in the relationship between teacher ratings of Attention-Deficit/Hyperactivity Disorder and observed classroom behavior. *J.Consult Clin.Psychol., 73,* 424-434.

32. Faraone, S. V., Perlis, R. H., Doyle, A. E., Smoller, J. W., Goralnick, J. J., Holmgren, M. A. et al. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57,* 1313-1323.

33. Faraone, S. V., Sergeant, J., Gillberg C., & Biederman, J. (2003). The worldwide prevalence of ADHD: Is it an American condition? World Psychiatry 2, 104-113.

34. Fombonne, E. (1994). The Chartres Study: I. Prevalence of psychiatric disorders among French school-age children. *Br.J.Psychiatry, 164,* 69-79.

35. Ford, T., Goodman, R., & Meltzer, H. (2003). The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV disorders. *J.Am.Acad.Child Adolesc.Psychiatry, 42,* 1203-1211.

36. Frick, P. J., Lahey, B. B., Applegate, B., Kerdyck, L., Ollendick, T., Hynd, G. et al. (1994). DSM-IV filed trials for the disruptive behavior disorders: Symptom utility estimates. *J.Am.Acad.Child Adolesc.Psychiatry, 33,* 529-539.

37. Gadow, K. D., Nolan, E. E., Litcher, L., Carlson, G. A., Panina, N., Golovakha, E. et al. (2000). Comparison of Attention-Deficit/Hyperactivity Disorder symptom subtypes in Ukranian schoolchildren. *J.Am.Acad.Child Adolesc.Psychiatry, 39,* 1520-1527.

38. Gallucci, F., Bird, H. R., Berardi, C., Gallai, V., Pfanner, P., & Weinberg, A. (1993a). Symptoms of attention-deficit hyperactivity disorder in an Italian school sample: findings of a pilot study. *Journal of the American Academy of Child & Adolescent Psychiatry, 32,* 1051-1059.

39. Gallucci, F., Bird, H. R., Berardi, C., Gallai, V., Pfanner, P., & Weinberg, A. (1993b). Symptoms of attention-deficit hyperactivity disorder in an Italian school

sample: findings of a pilot study. *Journal of the American Academy of Child & Adolescent Psychiatry, 32,* 1051-1059.

40. Gingerich, K. J., Turnock, P., Litfin, J. K., & Rosen, L. A. (1998). Diversity and attention deficit hyperactivity disorder. *J.Clin.Psychol., 54,* 415-426.

41. Graetz, B. W., Sawyer, M. G., Hazell, P. L., Arney, F., & Baghurst, P. (2001). Validity of DSM-IV ADHD subtypes in a nationally representative sample of Australian children and adolescents. *J.Am.Acad.Child Adolesc.Psychiatry, 40,* 1410-1417.

42. Halliburton, M. (2004). Finding a fit: Psychiatric pluralism in South India and its implications for WHO studies of mental disorder. Transcultural Psychiatry 41[1], 80-98.

43. Halperin, J. M., Matier, K., Bedi, G., Sharma, V., & Newcorn, J. H. (1992). Specificity of inattention, impulsivity, and hyperactivity to the diagnosis of attention-deficit hyperactivity disorder. *J.Am.Acad.Child Adolesc.Psychiatry, 31,* 190-196.

44. Halperin, J. M., Newcorn, J. H., Matier, K., Sharma, V., McKay, K. E., & Schwartz, S. (1993). Discriminant validity of attention-deficit hyperactivity disorder. *J.Am.Acad.Child Adolesc.Psychiatry, 32,* 1038-1043.

45. Hechtman, L. & Weiss, G. (1986). Controlled prospective fifteen year follow-up of hyperactives as adults: non-medical drug and alcohol use and anti-social behavior. *Canadian Journal of Psychiatry, 31,* 557-567.

46. Ho, D. Y. F. (1981). Childhood Psychopathology: A dialogue with special reference to Chinese and American cultures. In A.Kleinman & T. Lin (Eds.), *Normal and abnormal behavior in Chinese Culture* (pp. 137-155). Holland: Dordrecht.

47. Hopper, K. & Waderling, J. (2000). Revisiting the developed versus developing countr distinction in course and outcome in schizophrenia: Results from ISoS, the WHO collaborative project. Schizophrenia Bulletin 26[4], 835-846.

48. Johnston, C. & Mash, E. J. (2001). Families of children with attention-deficit/hyperactivity disorder: Review and recommendations for future research. *Clin.Child Fam.Psychol.Rev., 4,* 183-207.

49. Kleinman, A. (1988). *Rethinking psychiatry: From cultural category to personal experience.* New York: Free Press.

50. Lahey, B. B., Applegate, B., Barkley, R. A., Garfinkel, B., McBurnett, K., Kerdyck, L. et al. (1994a). DSM-IV field trials for Oppositional Defiant Disorder and Conduct Disorder in children and adolescents. *Am.J.Psychiatry, 151,* 1163-1171.

51. Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L. L., Hynd, G. et al. (1994b). DSM-IV trials for Attention-Deficit Hyperactivity Disorder in children and adolescents. *Am.J.Psychiatry, 151,* 1673-1685.

52. Leung, P. W., Luk, S. L., Ho, T. P., Taylor, E., Mak, F. L., & Bacon-Shone, J. (1996). The diagnosis and prevalence of hyperactivity in Chinese schoolboys. *Br.J.Psychiatry, 168,* 486-496.

53. Madras, B. K., Miller, G. M., & Fischman, A. J. (2005). The dopamine transporter and attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57,* 1397-1409.

54. Magnusson, P., Smari, J., Gretarsdottir, H., & Prandardottir, H. (1999). Attention-Deficit/Hyperactivity symptoms in Icelandic schoolchildren: Assessment with the Attention Deficit/Hyperactivity Rating Scale-IV. *Scand.J.Psychol., 40,* 301-306.

55. Mann, E. M., Ikeda, Y., Mueller, C. W., Takahashi, A., Tao, K. T., Humris, E. et al. (1992). Cross-cultural differences in rating hyperactive-disruptive behaviors in children. *Am.J.Psychiatry, 149,* 1539-1542.

56. Margalit, M. (1981). Cultural differences in the hyperactive syndrome rated in the Conners abbreviated scale. *J.Learn.Disabil., 14,* 330-331.

57. Marsella, A. & Kameoka, V. (1989). Ethnocultural issues in the assessment of psychopathology. In S.Wetzler (Ed.), *Measuring mental illness: Psychometric assessment for clinicians.* (pp. 231-256). Washington DC: American Psychiatric Association.

58. Matier-Sharma, K., Perachio, N., Newcorn, J., Sharma, V., & Halperin, J. (1995). Differential Diagnosis of ADHD: Are Objective Measures of Attention, Impulsivity, and Activity Level Helpful? *Child Neuropsychology, 1,* 118-127.

59. Millstein, R. B., Wilens, T. E., Biederman, J., & et al. (1997). Presenting ADHD symptoms and subtypes in clinically referred adults with ADHD. *Journal of Attention Disorders, 2,* 159-166.

60. Mitsis, E. M., McKay, K. E., Schulz, K. P., Newcorn, J. H., & Halperin, J. M. (2000). Parent-teacher concordance for DSM-IV attention-deficit/hyperactivity disorder in a clinic-referred sample. *J.Am.Acad.Child Adolesc.Psychiatry, 39,* 308-313.

61. MTA Cooperative Group (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. The MTA Cooperative Group. Multimodal Treatment Study of Children with ADHD. *Arch.Gen.Psychiatry, 56,* 1073-1086.

62. Nakao, K. & Treas, J. (1994). Updating occupational prestige and socioeconomic scores: How the new measures add up. Sociological Methodology 24, 1-72.

63. Nolan, E. E., Gadow, K. D., & Sprafkin, J. (2001). Teacher reports of DSM-IV ADHD, ODD, and CD symptoms in schoolchildren. *Journal of the American Academy of Child & Adolescent Psychiatry, 40,* 241-249.

64. O'Leary, K. D., Vivian, D., & Nisi, A. (1985a). Hyperactivity in Italy. *Journal of Abnormal Child Psychology, 13,* 485-500.

65. O'Leary, K. D., Vivian, D., & Nisi, A. (1985b). Hyperactivity in Italy. *Journal of Abnormal Child Psychology, 13,* 485-500.

66. Pineda, D., Ardila, A., Rosselli, M., Arias, B. E., Henao, G. C., Gomez, L. F. et al. (1999). Prevalence of attention-deficit/hyperactivity disorder symptoms in 4- to 17-year-old children in the general population. *Journal of Abnormal Child Psychology, 27,* 455-462.

67. Pineda, D. A., Lopera, F., Palacio, J. D., Ramirez, D., & Henao, G. C. (2003). Prevalence estimations of attention-deficit/hyperactivity disorder: differential diagnoses and comorbidities in a Colombian sample. *Int.J.Neurosci., 113,* 49-71.

68. Power, T., Andrews, T. J., Eiraldi, R., Doherty, B., Ikeda, M. J., & DuPaul, G. (1998). Evaluating Attention Deficit Hyperactivity Disorder Using Multiple Informants: The Incremental Utility of Combining Teacher With Parent Reports. *Psychological Assessment, 10,* 250-260.

69. Power, T. J., Costigan, T. E., Leff, S. S., Eiraldi, R. B., & Landau, S. (2001). Assessing ADHD across settings: contributions of behavioral assessment to categorical decision making. *J.Clin.Child Psychol., 30,* 399-412.

70. Prendergast, M., Taylor, E., Rapoport, J. L., Bartko, J., Donnelly, M., Zametkin, A. et al. (1988). The diagnosis of childhood hyperactivity. A U.S.-U.K. cross-national study of DSM-III and ICD-9. *J.Child Psychol.Psychiatry, 29,* 289-300.

71. Reichenbach, L. C., Halperin, J. M., Sharma, V., & Newcorn, J. (1992). Children's motor activity: reliability and relationship to attention and behavior. *Developmental Neuropsychology, 8,* 87-97.

72. Reid, R. (1995). Assessment of ADHD with culturally different groups: The use of behavioral rating scales. *School Psychology Review, 24,* 537-560.

73. Reid, R., Casat, C. D., Norton, J., Anastopoulos, A., & Temple, E. P. (2001). Using behavior rating scales for ADHD across ethnic groups: the IOWA Conners. *Journal of Emotional and Behavioral Disorders, 9,* 210-218.

74. Reid, R., DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Rogers-Adkinson, D., Noll, M. B. et al. (1998). Assessing culturally different students for attention deficit hyperactivity disorder using behavior rating scales. *Journal of Abnormal Child Psychology, 26,* 187-198.

75. Reid, R., Riccio, C., Kessler, R. H., DuPaul, G., Power, T., Anastopoulos, A. et al. (2000). Gender and ethnic differences in ADHD as assessed by behavior ratings. *Journal of Emotional and Behavioral Disorders, 8,* 38-48.

76. Reynolds, C. R. & Kamphaus, R. W. (2004). *Behavior Assessment System for Children- 2nd edition*. Circle Pines, Minnesota: AGS Publishing.

77. Rohde, L. A., Biederman, J., Busnello, E. A., Zimmerman, H., Schmitz, M., Martins, S. et al. (1999). ADHD in a school sample of Brazilian adolescents: A study of prevalence, comorbid conditions, and impairments. *J.Am.Acad.Child Adolesc.Psychiatry, 38,* 716-722.

78. Rowland, A. S., Lesesne, C. A., & Abramowitz, A. J. (2002). The epidemiology of attention-deficit/hyperactivity disorder (ADHD): a public health view. *Ment.Retard.Dev.Disabil.Res.Rev., 8,* 162-170.

79. Rutter, M. (1989). Isle of Wight revisited: twenty-five years of child psychiatric epidemiology. *J.Am.Acad.Child Adolesc.Psychiatry, 28,* 633-653.

80. Rutter, M., Tizard, J., Yule, W., Graham, P., & Whitmore, K. (1976). Research report: Isle of Wight Studies, 1964-1974. *Psychol.Med., 6,* 313-332.

81. Shaffer, D. (1994). Attention deficit hyperactivity disorder in adults. *American Journal of Psychiatry, 151,* 633-638.

82. Sonuga-Barke, E., Minocha, K., Taylor, E. A., & Sandberg, S. (1993). Inter-ethnic bias in teachers' ratings of childhood hyperactivity. *British Journal of Developmental Psychology, 11,* 187-200.

83. Still, G. (1902). Some abnormal psychical conditions in children. *Lancet, 1,* 1008-1168.

84. Swanson, J. & Carlson, C. L. (1994). DSM-IV rating scale for ADHD and ODD.

85. Tao, K. T. (1992). Hyperactivity and Attention Deficit Disorder syndromes in China. Journal of the American Academy of Child and Adolescent Psychiatry 31, 1165-1166.

86. Taylor, E. & Sandberg, S. (1984). Hyperactive behavior in English schoolchildren: a questionnaire survey. *Journal of Abnormal Child Psychology, 12,* 143-155.

87. Todd, R. D., Huang, H., Smalley, S. L., Nelson, S. F., Willcutt, E. G., Pennington, B. F. et al. (2005). Collaborative analysis of DRD4 and DAT genotypes in population-defined ADHD subtypes. *J.Child Psychol.Psychiatry, 46,* 1067-1073.

88. Tripp, G., Luk, S. L., Schaughency, E. A., & Singh, R. (1999). DSM-IV and ICD-10: a comparison of the correlates of ADHD and hyperkinetic disorder. *J.Am.Acad.Child Adolesc.Psychiatry, 38,* 156-164.

89. Wilens, T. E., Biederman, J., & Spencer, T. J. (2002). Attention deficit/hyperactivity disorder across the lifespan. *Annu.Rev.Med., 53,* 113-131.

90. Wolraich, M. L., Hannah, J. N., Baumgaertel, A., & Feurer, I. D. (1998). Examination of DSM-IV criteria for Attention Deficit Hyperactivity Disorder in a county-wide sample. *J.Dev.Behav.Pediatr., 19,* 162-168.

91. Wolraich, M. L., Hannah, J. N., Pinnock, T. Y., Baumgaertel, A., & Brown, J. (1996). Comparison of diagnostic criteria for Attention-Deficit/Hyperactivity Disorder in a county-wide sample. *J.Am.Acad.Child Adolesc.Psychiatry, 35,* 319-324.

92. Wolraich, M. L., Lambert, E. W., Baumgaertel, A., Garcia-Tornel, S., Feurer, I. D., Bickman, L. et al. (2003). Teachers' screening for attention deficit/hyperactivity disorder: comparing multinational samples on teacher ratings of ADHD. *Journal of Abnormal Child Psychology, 31,* 445-455.

93. Wolraich, M. L., Lambert, E. W., Bickman, L., Simmons, T., Doffing, M. A., & Worley, K. A. (2004). Assessing the impact of parent and teacher agreement on diagnosing attention-deficit hyperactivity disorder. *J.Dev.Behav.Pediatr., 25,* 41-47.

94. Yang, K. N., Schaller, J. L., & Parker, R. (2000). Factor structures of Taiwanese teachers' ratings of ADHD: A comparison with U.S. studies. *Journal of Learning Disabilities, 33,* 72-82.

95. Yao, K. N., Solanto, M. V., & Wender, E. H. (1988). Prevalence of hyperactivity among newly immigrated Chinese-American children. *J.Dev.Behav.Pediatr., 9,* 367-373.

96. Youngstrom, E., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *J Consult Clin.Psychol, 68,* 1038-1050.