

Diagnostic Accuracy and Clinical Potential for the Strengths and Weaknesses of
Attention-Deficit/Hyperactivity Symptoms and Normal Behaviors Scale (SWAN)

Ruqayah Ali Alhajji

A dissertation

submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Elaine Walsh, Chair

Monica Oxford

Mark Stein

Program Authorized to Offer Degree:

School of Nursing

© Copyright 2022

Ruqayah Ali Alhaji

University of Washington

Abstract

Diagnostic Accuracy and Clinical Potential for the Strengths and Weaknesses of Attention-Deficit/Hyperactivity Symptoms and Normal Behaviors Scale (SWAN)

Ruqayah Ali Alhajji

Chair of the Supervisory Committee:

Elaine Walsh

School of Nursing

Background: In theory, the continuum of attention-deficit hyperactivity disorder (ADHD) is normally distributed within the general population. Unlike many existing ADHD rating scales, the Strengths and Weaknesses of ADHD–Symptoms and Normal Behavior (SWAN) scale is theoretically designed to capture positive and negative variations in symptoms along this continuum. Nonetheless, despite its use in several research studies, the SWAN has received little attention to its clinical utility. Most studies that found the SWAN to have high validity and diagnostic accuracy for ADHD screening in children have been conducted in community/nonclinical samples. It is still unclear in clinical populations whether the SWAN captures a similar underlying construct relative to commonly used ADHD rating scales. There is inadequate information about the SWAN’s diagnostic validity compared with other validated ADHD scales to identify ADHD in clinic-referred children.

Aim: The aim of the current study was to estimate the SWAN’s convergent validity with the a common ADHD rating subscale, the Hyperactivity–Inattention (HI) of the Strengths and Difficulties Questionnaire (SDQ/HI). A second aim of the study was to compare the concurrent

validity of the SWAN with that of the SDQ/HI, as measured by correlations with scores of a measure of ADHD-related functional impairment, the Impairment Rating Scale (IRS). A third aim was to examine the diagnostic validity of the SWAN to differentiate between children with and without an ADHD diagnosis. A fourth aim was to compare the diagnostic validity of the SWAN and SDQ/HI to differentiate between children with and without an ADHD diagnosis.

Methods: This study used a cross-sectional cohort design to assess and compare the performance of two ADHD rating scales for identifying cases of ADHD. The study utilized archived data of 357 children who visited a mental health outpatient clinic at a children's hospital, specializing in ADHD and related disorders, in the northwest United States. One parent/caregiver completed the three rating scales, an average of 3 days before their child's first visit to the clinic for a comprehensive evaluation for ADHD. Bivariate correlation analyses were used to examine the convergent and concurrent validity among the current study rating scales. Receiver operating characteristic (ROC) curve analyses were used to determine the SWAN and SDQ/HI cutoff scores and compare their performance in identifying cases with ADHD from cases with other disorders in this sample.

Results: Regarding convergent validity, the SWAN total and SDQ/HI correlation was large and significant ($r = .68$). Likewise, the correlation between the SDQ/HI and the SWAN Hyperactivity/Impulsivity subscale (SWAN HI) was large and significant ($r = .65$). The correlation between the SDQ/HI and the SWAN Attention Deficit subscale (SWAN AD) was significant but smaller ($r = .49$). For concurrent validity, the correlation coefficients of the SWAN total and the SWAN subscales were larger than the correlation coefficient values with the IRS items compared to the SDQ/HI. Results from the ROC curves and Area Under the Curve (AUC) of the SWAN total, the SWAN AD and SWAN HI (AUCs = .7, .69 and .66, respectively)

showed good to useful abilities in differentiating between children with and without an ADHD diagnosis. The ROC curves analyses that compared the SWAN total and the SWAN HI subscale as well as the SWAN AD subscale with the SDQ/HI subscale when differentiating between the ADHD and non-ADHD groups indicated that the SWAN total had the largest AUC among all scales. However, the results from the pairwise comparison of the ROC curves with the critical value for significance for the AUC set at ($p = .05$) showed no significant differences among the AUCs of the SWAN total, the SWAN AD subscale, and the SDQ/HI. The discriminant abilities of the SWAN AD and the SWAN HI were slightly lower than the SWAN total. The difference between the SWAN total and the SWAN HI AUC was significant ($p = .023$).

Conclusion: The SWAN total and subscales showed good to useful abilities for classifying ADHD cases from non-ADHD clinic-referred children with other psychiatric disorders. While the SWAN demonstrated low specificity in the current study, these results might be specific to the current sample from an ADHD tertiary clinic, wherein the majority of referred cases were complex with significant psychiatric comorbidity, diverse medical and treatment history. Findings from comparing the diagnostic validity of the SWAN to SDQ/HI yielded no significant difference in their performance in identifying ADHD cases in the current sample. Nonetheless, the SWAN has advantages over SDQ/HI for describing aspects of variations (strengths) in attention and activity abilities.

Table of Contents

List of Figures iii

List of Tables iv

Chapter 1: Introduction 1

 Overview 1

 Problem Statement 4

 Study Purpose 5

 Study Aims & Hypotheses 6

Chapter 2: Review of Literature 8

Part I..... 8

 Development of Conceptualization of ADHD in the DSM 8

 Categorical Conceptualization of ADHD 9

 Shifting from a Categorical to a Dimensional Conceptualization of ADHD..... 10

 Proposed Dimensional Approaches to ADHD 11

 Modifications of the ADHD Diagnostic Criteria in DSM-5 13

 Integrating Dimensional Approach of ADHD in Clinical Practice 15

 Necessity of a Diagnostic Threshold 16

 Functional Impairment as an External Validator of the Disorder 17

 Decisions About Treatment 18

 Recognizing the Needs of Children with Subclinical ADHD 20

 Diagnostic Procedures for ADHD and Dimensional Rating Scales 21

 Complexity of ADHD Diagnostic Procedures..... 21

 Children with ADHD Symptoms and Challenging Clinical Profiles 22

Assessment of ADHD	23
Dimensional Rating Scales	23
Advantages & Limitations of ADHD Rating Scales	25
Part II	29
The Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale.....	29
Integrating the Continuum Approach to ADHD and the DSM Approach	29
Theoretical Statistical Cutoff	30
Strength-Based Assessment Approach of ADHD	31
The SWAN as a Research Instrument for ADHD Traits	32
Psychometric Properties of the SWAN	33
The SWAN Clinical Utility	33
The Convergent Validity of the SWAN with ADHD Symptoms Measures	34
The Concurrent Validity of the SWAN with Functional Impairment Measures	35
The Diagnostic Validity of the SWAN.....	36
Summary	40
Chapter 3: Methods	42
Study Design.....	42
Data Preparation.....	42
Participants	43
Measures	46
Definition of the ADHD Criterion Diagnosis	46
The Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale.....	47
Strengths and Difficulties Questionnaire (SDQ)	48

The Impairment Rating Scale (IRS)	49
Data Analysis	50
Chapter 4: Results	55
Sample.....	56
Convergent Validity Findings	57
Concurrent Validity Findings	57
Diagnostic Validity Findings	59
Comparing Diagnostic Validity of SWAN with SDQ/HI Findings.....	60
Chapter 5: Discussion and Conclusions	62
Convergent Validity.....	63
Concurrent Validity	63
Diagnostic Validity	66
Comparing Diagnostic Validity of SWAN with SDQ/HI	68
Limitations and Strengths	70
Future Research Direction	71
Clinical Implications	72
Conclusion	74

List of Figures

Figure 3.1 Flow chart for the data preparation and selection of the study participants based on the study inclusion criteria	93
Figure 4.1. Correlations between the SWAN total and subscales with SDQ Hyperactivity-Inattention subscale.....	94
Figure 4.2. Receiver operating characteristic (ROC) curves indicating discriminant abilities of the SWAN Total, SWAN, HI, SWAN AD in identifying ADHD group relative to non-ADHD group.....	95
Figure 4.3. The Receiver operating characteristic (ROC) curve indicating discriminant abilities of the Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale in identifying ADHD group relative to non-ADHD group.....	96
Figure 4.4. Receiver operating characteristic (ROC) curves comparing discriminant abilities among SWAN Total, SWAN, HI, SWAN AD, and SDQ HI for identifying ADHD group relative to non-ADHD group.....	97

List of Tables

Table.2.1. The Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) Symptoms of ADHD.....	98
Table 3.1. Demographic and Clinical Characteristics of Children in ADHD and Non-ADHD groups.....	99
Table 4.1 Descriptive Statistics for Each of the Study Scales.....	101
Table 4.2. Bivariate Correlations of Convergent Validity of the SWAN total and subscales with SDQ Hyperactivity-Inattention subscale.....	102
Table 4.3. Bivariate Correlations of Concurrent Validity presenting correlations between IRS items, SWAN total and subscales with SDQ Hyperactivity-Inattention.....	103
Table 4.4. Areas Under the Curve, Sensitivities, Specificities, and Optimum Cutoff Scores for the SWAN Total, SWAN, HI, SWAN AD, and SDQ HI.....	104

Acknowledgements

The completion of my Ph.D. studies would not have been possible without the help of so many people who loaded me with love and support—to you all, I am grateful.

First and foremost, I wish to thank my dissertation committee members. I am grateful and privileged to work with truly remarkable scientists and mentors who provided me with their time, expertise, and support during this journey.

I would like to sincerely acknowledge my committee chairperson, Dr. Elaine Walsh, for her invaluable guidance, support, and patience during the course of my Ph.D. I have been fortunate to have a supervisor who has been committed to my successes and was genuinely supportive during the ups and downs of my journey. Dr. Walsh was always there when I needed her feedback and guidance and promptly responded to my questions and queries. Dr. Walsh celebrated every small work progress I accomplished and always reminded me to take care of myself.

I would also like to express my gratitude to Dr. Mark Stein, who provided me with professional motivation and insight into a complex field. This project would be impossible without Dr. Stein's generosity, guidance, and support. Thank you, Dr. Stein, for so many opportunities you offered and your belief in my passion for helping children with ADHD and their families. I look for you as a mentor and role model for my professional carrier. I want to extend my appreciation to Dr. Monica Oxford. Dr. Oxford's insightful feedback pushed me to sharpen my writing continuously and brought my work to a higher level. Special thanks to Dr. Ken Pike, whose statistical expertise and suggestions guided me throughout this project. Further, I would like to thank Dr. Freda Liu for her efforts and contribution in obtaining the data for this study that ultimately made this work possible.

I want to thank my parents for their unconditional love and unwavering support. Mom, you have always been a true inspiration for so many women to follow their passion, pursue their dreams, and serve their community. It is such an honor to be your daughter and to have my mother as my role model. Dad, your determination, grit, and preservations have always motivated me to continue toward my goals, whatever comes my way. To my siblings: Sadiq, thank you for taking care of the kids whenever you were around, so I had to finish my work. Hassan, thank you for always sending me beautiful pictures of my sweetest nieces, who I cannot wait to spend more time with.

To my husband Hassan, without whom I would never have the courage to take on such a challenging journey. Thank you for sacrificing so much to provide me with the opportunity to achieve this dream. Thank you as many as those thousands of miles and hundreds of hours you traveled back and forth over the past years between Riyadh and Seattle. Thank you; as much as those endless times I have told you I can't do this, it is so hard, and you always responded: I know you can do it. Special thanks to my sons Jawad and Fears, my biggest cheerleaders and partners in this journey that would never be the same without you. Thank you for your patience when I had to work so many hours and not be available all the time. I cannot believe how fast you grew. I am proud of how intellectual, responsible, funny, and compassionate young men you are today.

I am thankful for the unlimited support of so many friends I met here in Seattle. Thank you so much for making so many of Seattle's gloomy gray days more shiner. Finally, I would like to express my sincere appreciation for King Saud University and the Saudi Arabian Cultural Mission for their financial support. I was very grateful that I received the Hester McLaws nursing scholarship that helped me complete my Ph.D. project.

Chapter 1: Introduction

The purpose of this chapter is to briefly introduce a study that focuses on the validity and clinical utility of a rating scale of attention-deficit/hyperactivity disorder (ADHD)—the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale (SWAN)—in a sample of clinically referred children aged 6–11 years. The study problem statement, purpose, aims, and hypotheses will then be summarized.

Overview

Attention-deficit hyperactivity disorder (ADHD) is one of the most common neurodevelopmental conditions of childhood. In 2016, in a nationally representative survey of parents' reports regarding their children's health, 2.4 million children aged 6–11 years in the United States had an ADHD diagnosis (Danielson et al., 2018). ADHD is characterized by impairing symptoms of inattention, hyperactivity, and impulsivity (American Psychiatric Association, 2013). Core ADHD symptoms interfere with many aspects of a child's daily functioning, such as academic performance, social skills, and parent-child relations (Garner et al., 2013a; Kuriyan et al., 2013). Evidence-based treatments help with managing ADHD symptoms (Barbareisi, 2020; Pliszka, 2007; Wolraich et al., 2019). An accurate diagnosis is the first step toward effective treatment. However, ADHD diagnosis is a complex process with numerous challenges (Parker & Corkum, 2016). The clinical picture of children with ADHD is complicated because ADHD is often associated with various comorbidities. Many medical and psychiatric disorders symptoms are similar to symptoms manifest in primary ADHD (Pham & Riviere, 2015; Pliszka, 2003).

Most ADHD guidelines recommend applying rating scales to gather information from different informants (e.g., parent, teacher, and caregiver) prior to diagnosis (Pliszka, 2007;

Wolraich et al., 2019). Although rating scales are not considered sufficient for clinical ADHD diagnosis, they are endorsed as good clinical practice for ADHD screening purposes, measuring symptom severity, and treatment monitoring (Collett, Ohan, & Myers, 2003; Pliszka, 2007).

Evidence supports a dimensional conceptualization for ADHD (Coghill & Sonuga-Barke, 2012; Faraone et al., 2015). This conceptualization suggests that ADHD reflects the extreme end of a range of behavioral dimensions of attention and activity control and is distributed in the population on a continuum along which individuals differ by degree of severity rather than in kind (Coghill & Sonuga-Barke, 2012; Swanson, et al., 2012). Using behavioral rating scales is a typical method that provides dimensional supplements to ADHD clinical diagnoses by quantifying symptom severity (Swanson, Wigal, & Lakes, 2009). ADHD rating scales, including both narrowband (i.e., ADHD-specific items developed based on the ADHD symptoms from the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.) (DSM-5) (American Psychiatric Association, 2013), and broadband (i.e., assessing various behavioral conditions related to general psychopathology, along with ADHD-related symptoms) yield factor or subscale scores with cutoff points that can be used to assist in a categorical diagnosis that defines ADHD dichotomously as present or absent (Collett et al., 2003).

However, these scales can be described as unidirectional and focusing on symptom severity rather than full-continuum behavioral rating scales: they do not identify a child's strengths on a particular dimension (Alexander, Salum, Swanson, & Milham, 2020b; Greven et al., 2016). They only measure moderate to severe symptoms at one end of the ADHD distribution and ignore existing differences among less symptomatic and normally functioning individuals in the general population. Thus, most of the summary score results of these rating scales are likely to produce truncated distributions due to discontinuity of measuring variations

in ADHD dimensions beyond the diagnostic threshold in population samples. Despite that, most common rating scales use linear standardizations of the raw scores to set cutoff points derived from population norms to define moderate to severe cases of ADHD that require clinical attention. Consequently, these rating scales share the same risk because when the normal distribution assumption is violated, it is problematic to apply such statistical cutoffs (Swanson et al., 2012).

With the goal to overcome limitations of the current ADHD rating scales, Swanson et al. (2012) developed the SWAN rating scale to identify individuals with strengths (e.g., above-average attention and activity control skills) and those with weaknesses (e.g., inattention and hyperactivity-impulsivity that indicate a clinical ADHD diagnosis) along ADHD behavioral dimensions. The narrowband approach was also considered to ensure agreements between the SWAN's scores and ADHD clinical diagnosis based on *DSM* criteria. Each item of the SWAN represents each *DSM* ADHD symptom converted from negative statements (e.g., does not seem to listen when spoken to directly) to more neutral statements (e.g., listens when spoken to directly). As a result of rephrasing, each item of the SWAN is a bidirectional statement that could capture variations along the underlying ADHD dimensions of behavior at every item level.

This neutral language of the SWAN might minimize the stigma associated with reporting only a child's problems by allowing the informants to report the child's strengths, if any exist (Polderman et al., 2007; Robaey, Amre, Schachar, & Simard, 2007). A similar approach has already been implemented by the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997; Swanson, Wigal, Lakes, & Volkow, 2012). The SDQ is a broad band rating scale that assesses various behavior and emotional problems associated with many psychopathological disorders involving ADHD. Some SDQ items assess strength (e.g., sees tasks through to the end,

and good attention), and others assess problems (e.g., constantly fidgeting or squirming).

However, while the SDQ allows the informants to report child strengths only in some aspects of a child's behavior, the SWAN allows rating strengths and problems simultaneously for all *DSM-5* ADHD symptoms.

The SWAN has the advantage of increasing the statistical power needed for research studies because it measures variations of ADHD dimensions even among normally functioning individuals. Consequently, the SWAN has been used by various genetic studies (Burton et al., 2019; Hay, Bennett, Levy, Sergeant, & Swanson, 2007; Polderman et al., 2007). It has also been used to examine some neuropsychological aspects of ADHD and assess response to medical stimulants and other trials for ADHD interventions (Bosch et al., 2020; Siebelink et al., 2021; Wang et al., 2021). Nonetheless, despite its use in several research studies, the SWAN has received little attention regarding its clinical utility.

Problem Statement

Several studies examined the SWAN psychometric properties. Results from studies that examined the SWAN's convergent validity, which refers to the degree of correlation between two scales that measure similar constructs, showed that the SWAN had high correlations with narrowband scales and low to moderate correlations with broadband scales. However, most of these studies have been conducted in community samples (Lakes, Swanson, & Riggs, 2011; Polderman et al., 2007). Thus, there is a lack of information regarding whether the SWAN collects information similar to other validated and widely used rating scales in the clinic population. The studies that have explored the SWAN's diagnostic accuracy were mainly based on discriminating between children with ADHD and healthy controls (Burton et al., 2019; Lai et al., 2011; Schulz-Zhecheva et al., 2017). Typically, employing a sample of healthy controls

yields less clinically applicable findings because clinicians are most likely to decide whether a child has ADHD, some other diagnosis, or comorbid diagnoses (Raiker et al., 2017).

As yet, only two studies (French and Chinese translated forms) examined the SWAN's diagnostic validity among clinical samples of children to differentiate ADHD from other psychiatric diagnoses (non-ADHD) (Chan, Lai, Luk, Hung, & Leung, 2014; Robaey et al., 2007). These two studies provided preliminary evidence that the SWAN demonstrated good to excellent abilities to discriminate between ADHD and other disorders in clinical settings. However, both studies used parent-structured interviews as references for ADHD diagnosis, which might increase the risk of shared method variance between the SWAN scores and clinical diagnosis, as both were obtained from parents. Therefore, results from these studies might also overestimate the SWAN's usefulness in clinical settings. To better estimate the SWAN diagnostic accuracy in a clinical population, a reference standard based on a best-practices comprehensive assessment that integrates information from multiple sources (e.g., semi-structured interviews with the parent and the child, developmental history, and teacher rating scales) should be used. To date, no study has evaluated the SWAN-English's diagnostic validity for the SWAN total and subscales for a clinic-referred sample, who typically are at risk for high levels of impairments and comorbidities compared to their community counterparts, based on the best-practice comprehensive assessment for ADHD diagnosis.

Study Purpose

The purpose of this study was to address current gaps in the literature about the SWAN's applicability in clinical settings. Additionally, this study compared the SWAN's psychometric properties (convergent, concurrent, and diagnostic validity) with a common ADHD rating subscale, the SDQ hyperactivity-inattention subscale (SDQ/HI), in a sample of 357 clinically

referred children aged 6–11 years. Children were referred based on suspicion of their having attentional and behavioral problems. All participants received a comprehensive clinical evaluation to determine their diagnoses. One parent/caregiver completed the rating scales included as part of a pre-visit online screening routine, either on the same day or within a short time period, on average 3 days before their child's first visit to the clinic.

Study Aims and Hypotheses

The specific aims of this study were to:

- 1) Estimate the SWAN's convergent validity with the SDQ/HI by examining the SWAN total and subscale scores correlations with the SDQ/HI

Hypothesis 1. The correlations between the SWAN total and subscales scores and SDQ/HI subscales scores will be both significant and moderate, as evidenced by correlation coefficient equal to or greater than 0.5.

- 2) Compare the concurrent validity of the SWAN with that of the SDQ/HI, as measured by correlations with scores of a measure of ADHD-related functional impairment, the Impairment Rating Scale (IRS)

Hypothesis 2. The SWAN total will have larger correlations with the IRS items than with the SDQ/HI, as evidenced by the significant and moderate correlation coefficient equal to or greater than 0.5.

- 3) Examine the diagnostic validity of the SWAN by exploring the following:
 - a) The discriminant abilities of the SWAN total and subscales scores to differentiate between children with and without an ADHD diagnosis
 - b) The diagnostic efficiency of the SWAN total and subscales scores to determine the best cutoff scores that differentiate between children with and without an ADHD diagnosis

Hypothesis 3. The SWAN total and subscales will demonstrate good discriminant abilities to discriminate between children with and without ADHD.

- 4) Compare the diagnostic validity of the SWAN total and subscales and SDQ/HI scores to differentiate between children with and without an ADHD diagnosis

Hypothesis 4. The SWAN total scores will have greater discriminant ability to differentiate between children with and without an ADHD diagnosis than the SDQ/HI.

Chapter 2: Review of Literature

Introduction

The purpose of this chapter is to provide a review of the literature pertinent to the key elements of the study. The first section highlights how advances in ADHD conceptualization influenced the ADHD definition in the nosology as outlined in different versions of the *Diagnostic and Statistical Manual for Mental Disorders*, currently in its fifth edition (*DSM-5*; APA, 2013). This section includes a discussion about the categorical and dimensional approaches and conceptualizations of ADHD and their advantages and limitations. This will be followed by examining the implications of using ADHD behavioral rating scales as a dimensional complement to the ADHD categorical approach, with particular emphasis on the consequences of their psychometric limitations. The second section of this review focuses on the validity of an ADHD behavioral rating scale, the SWAN scale, which stands to make important contributions to correct shortcomings of most common ADHD rating scales. Nonetheless, despite its use in several research studies, the SWAN has received little attention to its clinical utility. The overarching focus of this study is on the SWAN's clinical utility. Thus, I will summarize the literature about current best practices of assessment of clinical diagnosis of ADHD and state of the field about the concurrent, convergent, and diagnostic validity of the SWAN.

Part I:

Development of Conceptualization of ADHD in the *DSM*

Historically, the conceptualization of ADHD has advanced significantly from an early description as “minimal brain dysfunction” (Lange, Reichl, Lange, Tucha, & Tucha, 2010) to its current definition in *DSM-5* as a “persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development” (APA, 2013, p. 59). The first

contemporary description of ADHD was in the *Diagnostic and Statistical Manual of Mental Disorders* (2nd ed.) (*DSM-2*), published in 1968, which referred to behaviors as a hyperkinetic reaction of childhood (Lange et al., 2010). However, it was not until *DSM-3*, published in 1980, that the disorder was recognized as attention deficit disorder (ADD) (with or without hyperactivity) (Lange et al., 2010).

Categorical Conceptualization of ADHD

DSM-3 marked a major shift in modern psychiatric classification. It introduced significant changes that increased the reliability of psychiatric diagnoses (Joober & Tabbane, 2019). *DSM-3* adopted a *categorical* approach to classification supported by the neo-Kraepelinian work which attempted to shift psychiatry closer to a traditional medical model (Joober & Tabbane, 2019; Sonuga-Barke, 1998). This medical categorical approach treated childhood disorders (e.g., ADHD) as distinct entities, with each disorder having its own clearly defined diagnostic criteria (Sonuga-Barke, 1998). More importantly, it provided operationalized criteria to diagnose psychiatric disorders, thereby decreasing clinical uncertainty and diagnostic variability (First, 2010). Further, *DSM-3* was intended to identify homogeneous subgroups of individuals who probably shared a common underlying etiology. For this reason, *DSM-3* defined several diagnostic subtypes and divided broad diagnostic groups into narrowly defined categories that shared specific diagnostic features (First, 2010). *DSM-3*'s categorical approach was widely accepted by clinicians because it facilitated better communication among professionals and informed clinical evaluations, planning, and treatments (Clark, Cuthbert, Lewis-Fernández, Narrow, & Reed, 2017).

DSM-3 re-conceptualized ADHD into two sub-diagnostic categories, ADD with and ADD without hyperactivity, that reflected the categorical approach adopted by *DSM-3* (Frick &

Nigg, 2012). The two sub-diagnostic categories were created because it had been unclear whether the “ADD subtype attention deficit without hyperactivity” was qualitatively similar to that of the “ADD subtype with hyperactivity” or whether the two forms should be regarded as two different clinical conditions (Frick & Nigg, 2012; Lange et al., 2010). However, with the publication of the revision of *DSM-3* in 1987 (*DSM-3-R*), the symptoms of inattention, impulsivity, and hyperactivity were pooled together into one list and conceptualized as a single construct with a single cutoff score, and this category remains to this day: ADHD. In *DSM-4* (APA, 1994), and *DSM-IV-TR* (APA, 2000), although ADHD retained its name, ADHD subtypes were reintroduced as inattentive, hyperactive/impulsive, combined, and not otherwise specified.

Shifting from a Categorical to a Dimensional Conceptualization of ADHD

Prior to the release of *DSM-5*, the problems and shortcomings of the categorical conceptualization of ADHD presented in previous *DSM* versions became clear (Brown & Barlow, 2005; Coghill & Sonuga-Barke, 2012; Frick & Nigg, 2012; Rohde, 2008; Swanson, J. M. et al., 2012). The categorical approach to ADHD has been criticized because it separates the normal from the pathological by kind rather than by degree, as if a disorder is either present or absent in an individual. This approach adopted an “essentialist” perspective as if psychiatric disorders (e.g., ADHD) are “essence-based” or “natural kinds,” characteristically indefinable by fixed and inherent biological properties, meaning that individuals with the disorder qualitatively differ from those without the disorder.

However, recent conceptualizations of ADHD are moving away from a categorical view that defines ADHD as a unitary entity that is either present or absent in a given individual (Faraone et al., 2015; Marcus & Barry, 2011). Researchers in the field have been more inclined toward a *dimensional* approach to ADHD due to cumulative evidence supporting this view

(Greven, Buitelaar, & Salum, 2018; Hay et al., 2007; Polderman et al., 2007; Swanson, J. M. et al., 2012). This dimensional view indicates that children with ADHD differ from those without ADHD by degree of severity rather than in kind (Coghill & Sonuga-Barke, 2012). Evidence that supports this view has accumulated from several studies, including behavioral genetics (Burton et al., 2019; Faraone & Larsson, 2019), taxonomic studies (Frazier, Youngstrom, & Naugle, 2007; Haslam et al., 2006), and neurocognitive studies (Salum et al., 2014). Evidence suggests that ADHD reflects the extreme end of a range of behavioral dimensions of attention and activity control and is distributed in the population on a continuum along which individuals differ only by degree (Coghill & Sonuga-Barke, 2012; Swanson et al., 2012).

The dimensional approach to conceptualizing ADHD has various advantages that include allowing clinicians and researchers to quantify ADHD severity and evaluate changes in symptoms over time. Using behavioral rating scales is a typical method that provided dimensional supplements to ADHD clinical diagnoses by quantifying symptoms severity and monitoring treatment response (Swanson et al., 2009). The benchmark study of the National Institute of Mental Health (NIMH): the Multimodal Treatment of Attention Deficit Hyperactivity Disorder (MTA), a 14-month randomized clinical trial, used the rating scales as complements to the categorical diagnosis of ADHD based on a semi structured parent interview to evaluate changes in symptoms over time (Murray et al., 2008). More importantly, the dimensional approach to ADHD has facilitated some scientific discoveries, such as the genetics of problems with inattention or hyperactivity/impulsivity (Frazier et al., 2007; Salum et al., 2014).

Proposed Dimensional Approaches to ADHD

The debate about dimensional versus categorical assessment and conceptualizations of child psychopathology spans many decades (Achenbach, 1966; Marcus & Barry, 2011; Shekim

et al., 1986; Sonuga-Barke, 1998). However, prior to the publication of *DSM-5* there was evidence supporting the idea that most psychiatric disorders (e.g., ADHD) are better defined dimensionally. Thus, there was a strong interest in exploring the feasibility of alternative dimensional approaches for the different psychiatric diagnostic categories (First, 2010), including ADHD (Frick & Nigg, 2012; Marcus & Barry, 2011; Swanson et al., 2009).

Coghill and Sonuga-Barke (2012) provided an in-depth analysis of the philosophical and empirical debate on applying categorical or dimensional approaches to child and adolescent mental disorders. They identified three proposed dimensional approaches in the literature to conceptualize ADHD. The first was based on studies of ADHD's underlying structure, which suggests, based on confirmatory factor analysis, that the best ADHD model fit is a bifactor model with a general factor and specific factors of inattention and hyperactivity–impulsivity. Thus, this approach suggested replacing the ADHD category with a description of an individual's scores on symptom-rating measures on each of ADHD's dimensions (inattention, hyperactivity–impulsivity) and computing a total ADHD score.

The second proposed dimensional approach to ADHD involved advocating for the replacement of assessment based on current diagnostic categories of the *DSM* with empirically derived dimensions created by applying multivariate statistical analyses that capture a wide range of children's problematic behaviors. Prominent examples of this approach are the Child Behavior Checklist (CBCL) (Achenbach, 1991) and the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). While these rating scales contain some subscales that are not necessarily related to a specific *DSM* condition (e.g., CBCL has a sleep problem subscale, and SDQ has a prosocial score), they also contain subscales that tap into *DSM* categories (e.g., CBCL—subscale of an attention problem, and SDQ—subscale hyperactivity symptoms).

The third approach suggested combining the categorical and dimensional models, which has been promoted as a balanced and preliminary step toward adapting the dimensional model of ADHD to clinical practice. This approach does not abandon the *DSM* categorical model but instead aims to integrate dimensional approaches into current categorical clinical practice (Coghill & Sonuga-Barke, 2012; Frick & Nigg, 2012; Swanson et al., 2009).

Modifications of the ADHD Diagnostic Criteria in DSM-5

Despite the strong argument to include a dimensional component in the updated ADHD description in *DSM-5* (Frick & Nigg, 2012; Marcus & Barry, 2011; Swanson et al., 2009), the dimensional definition of ADHD is not addressed in *DSM-5* (Polanczyk, 2014). Although different dimensional approaches have been proposed to redefine ADHD, none of them provided a compelling alternative model with equivalent rigor and clinical applicability as the approach that was outlined in *DSM-IV* (Posner, Polanczyk, & Sonuga-Barke, 2020). *DSM-5* retains the previous *DSM-IV* ADHD symptom domains and its 18 core symptoms: nine involving inattention and nine involving hyperactivity–impulsivity (see Table-1) (APA, 2013). This preservation of most of *DSM-IV*'s diagnostic criteria was based on the evidence from extensive research that these criteria had stability and validity over time in identifying individuals with significant and persistent impairment. It also suggests that the cumulative evidence based on these *DSM-IV* criteria can be generalized in future studies (Willcutt et al., 2012).

Nonetheless, *DSM-5* made some modest modifications of the ADHD diagnostic criteria that reflected the recognition of the dimensional nature of the condition. In *DSM-5*, ADHD categorical diagnosis has been reconceptualized from merely binary diagnosis into an ordinal classification with three ordered values of severity (mild, moderate, and severe) based on the

clinically relevant information such as symptom counts, duration, and degree of impairment (Epstein & Loren, 2013).

In contrast to *DSM-IV*, *DSM-5* does not maintain the three ADHD subtype categorizations. In *DSM-5*, ADHD symptoms are divided into three presentation sets: ADHD inattentive presentation, ADHD hyperactivity–impulsivity presentation, and ADHD combined presentation (both inattentive and hyperactivity–impulsivity symptoms). Although *DSM-5*'s three presentations may seem similar to the *DSM-IV* subtypes, the labeling shift was more consistent with the ADHD dimensional model of severity. This labeling implied a departure from “subtypes”—which suggest more unchanging, stable traits—in favor of “presentations,” which reflect current evidence on ADHD symptom profiles, and which recognizes that individuals can alternate between ADHD symptoms “presentations” across their lifetime.

Multiple studies had failed to provide evidence for the validity of the ADHD subtypes provided in *DSM-IV* (Frick & Nigg, 2012; Willcutt et al., 2012). These studies demonstrated that ADHD subtypes had temporal instability, and that inconsistent symptom counts were reported by different informants, caregivers, and teachers (Valo & Tannock, 2010). The hyperactivity–impulsivity subtype was infrequent except among preschool samples and has been considered a developmental antecedent to the combined subtype (Frick & Nigg, 2012; Willcutt et al., 2012).

Some argued that there was supporting evidence of three valid subtypes given that the level of impairment differs among the individuals in the three subtypes, with combined subtypes having the most difficulty (Hurtig et al., 2007). However, some challenged this conclusion and stated that this instead provided evidence that ADHD represents the end of the spectrum of two continuous dimensions because children with combined ADHD have more symptoms (Nigg, 2015). Thus, as Nigg (2015) explained, these subtypes have been arbitrarily divided on the

ADHD spectrum into subtypes of severity instead of distinct valid subtypes. Nevertheless, the reasons that *DSM-5* did not completely abandon the subtypes and only slightly changed the labeling are to acknowledge the heterogeneity of ADHD and the lack of an alternative subtyping model (Nigg, 2015).

An additional modification in ADHD diagnostic criteria in *DSM-5* is the definition of impairment. Unlike *DSM-IV*, which only considered clinically significant symptoms in cases of severe impairments, *DSM-5* only requires evidence that the ADHD symptoms interfere with a child's functioning. This modification might indeed increase diagnostic reliability given that symptoms are a more objective criterion than impairments, which is a more subjective criterion (Epstein & Loren, 2013).

Integrating Dimensional Approach of ADHD in Clinical Practice

All evidence supports a dimensional definition of ADHD (Faraone et al., 2015). Yet, that does not mean abandoning the categorical approach to ADHD entirely (Coghill & Sonuga-Barke, 2012; Frick & Nigg, 2012; Marcus & Barry, 2011; Van Os et al., 1999). Both approaches are equally important. An integrated approach is the goal by incorporating dimensional models of ADHD over time to advance the categorical (nosological) diagnostic criteria and increase applicability of this integrated approach for clinical decisions and research (Coghill & Sonuga-Barke, 2012; Frick & Nigg, 2012; Lacalle, Ezpeleta, & Doménech, 2012; Swanson, et al., 2012; Swanson et al., 2009). Adopting the ADHD dimensional approach in clinical practice is challenging. The dimensional approach entails practices that include emphasizing levels of impairment instead of diagnostic thresholds, applying dimensional assessment, and assigning different treatments based on the measurable differences between individuals on the ADHD continuum (Haslam et al., 2006). Some have argued the clinical utility of the dimensional

approach to ADHD given that the clinical decision is categorical in nature. Nevertheless, clinicians can separate children who need treatment from those who do not and can select the most appropriate treatment without discarding the theoretical bases that these children differ by degree rather than kind (Coghill & Sonuga-Barke, 2012). Thus, there is no conflict in holding a dimensional understanding of ADHD and recognizing the clinical necessity of making a practical decision about treatment.

Necessity of a Diagnostic Threshold

According to Frick and Nigg (2012), integrated approaches to ADHD at the very least should acknowledge the clinical necessity of a diagnostic threshold. Thus, while assuming ADHD is best described and measured dimensionally along a continuum, there is still a pragmatic clinical need to impose a clear demarcation, “a diagnostic threshold” on such a continuum, to decide about treatment (Coghill & Sonuga-Barke, 2012; Haslam, 2002).

Advocates for integrating a dimensional approach of ADHD into clinical practice suggest that ADHD behavioral dimensions are analogous to many key health indicators, such as blood pressure, cholesterol, body mass index, and glucose levels (Asherson & Trzaskowski, 2015). All these indicators fall on continuums with normal variation within the general population, with no clear-cut biological or natural distinction between normal and abnormal. However, elevation of such indicators at some point along the continuums is associated with greater risk of hypertension, myocardial infarction, obesity, and diabetes (Haslam, 2002). Thus, such conditions (e.g., ADHD) fit better with the “*altered function model*,” which could represent any condition that impairs functioning and poses health risks because they are more likely to predict long-term suffering, impairments, or mortality (Zachar & Kendler, 2007).

Functional Impairment as an External Validator of the Disorder

From a dimensional view of ADHD, symptoms can exist in the absence of impairments (Epstein & Loren, 2013; Robaey et al., 2007). Some have questioned whether there is a true placement of a threshold to create a discontinuity along the ADHD continuum or whether this threshold is just arbitrarily assigned. For this reason, it has been recognized that it is crucial to rely on functional impairment as an external validator to establish a diagnostic threshold (Rohde, 2008). Functional impairment is a common cause for clinical referral (Sasser, Schoenfelder, & Stein, 2017), often with parents seeking treatment to help their children to overcome daily challenges (Weiss, McBride, Craig, & Jensen, 2018). A diagnosis of ADHD requires both presence of symptoms and specific impairment that interferes with daily life functioning (APA, 2013). Thus, when impairment is subclinical or due to other causes (e.g., learning disorder, depression, anxiety, and poor conduct), some children with ADHD behavioral symptoms may not meet ADHD diagnostic criteria.

Critics contend that shifting the focus of the ADHD criteria in *DSM-5* from impairment to symptoms might increase ADHD prevalence rates when taking into account that ADHD symptoms might appear even if there is no impairment (Epstein & Loren, 2013). Although ADHD symptoms and functional impairment seem intertwined, they are independent constructs (Sasser et al., 2017). ADHD symptoms represent the behavioral manifestations of ADHD; functional impairments are considered the consequences of these manifestations (Weiss et al., 2018). Symptoms of ADHD lead to complex functional and behavioral challenges in several life domains and usually require a management plan with multifold treatment targets and interventions. There is substantial evidence that ADHD is associated with adverse outcomes for children with the disorder, their families, and society overall. In educational settings, children's

inattention symptoms (e.g., distractibility, off-task behavior, not following instructions, and disorganization) can cause academic underachievement and interpersonal difficulties with peers and teachers (Garner et al., 2013b; Kuriyan et al., 2013; Swanson et al., 2017). Compared to their counterparts without ADHD, children with ADHD experience poorer social relationships, more peer rejections, impaired friendships, and more social performance difficulties, likely due to inattention and impulsive and increased activity behaviors (Schatz et al., 2020). In the family context, parents of children with ADHD experience high parenting stress, impaired parent–child relationships, and high rates of marital conflict (Franke et al., 2018; Kuriyan et al., 2013; Schatz et al., 2020; Swanson et al., 2017).

In general, part of the ADHD evaluation is to ask parents to report a child’s degrees of impairment in various contexts (e.g., home, school, and social interactions with peers). This information is incorporated into the diagnostic-decision process and treatment plan. A comprehensive functional impairment assessment that addresses lack of skills and functional deficits is crucial for making an accurate diagnosis, creating a treatment plan, and sequencing interventions (Sasser et al., 2017).

Decisions About Treatment

Acknowledging the underlying structure of ADHD as a continuum does not change the need for evidence-based treatment for children at the problematic end of the spectrum (McLennan, 2016). If not treated, children with ADHD are at greater risk for longer-term adverse outcomes in adolescence and adulthood (Franke et al., 2018), such as lower educational and employment attainment (Kuriyan et al., 2013), more traffic accidents (Chang, Lichtenstein, D’Onofrio, Sjölander, & Larsson, 2014) and higher rates of substance abuse (Groenman, Janssen, & Oosterlaan, 2017).

Currently, there is agreement among all guidelines to recommend multimodal treatment for ADHD, which is defined as combining two evidence-based modalities, such as psychosocial interventions and pharmacological treatments, to manage ADHD symptoms and functional impairments (Barbarese, 2020; Pliszka, 2007). Effectiveness of multimodal treatment of ADHD is supported by compelling evidence provided by the Multimodal Treatment of Attention Deficit Hyperactivity Disorder (MTA) study that compared four treatment arms: medication management, behavior modification, the combination of the first two, or treatment-as-usual for a community comparison (Murray et al., 2008). The effects of multimodal therapy (methylphenidate and intensive behavioral therapy) were significantly more effective after 14 months than behavioral intervention only or than “standard” therapy (treatment-as-usual in the community) of the control group (Arnold, Hodgkins, Caci, Kahle, & Young, 2015). Extensive evidence has supported the use of long-acting stimulant medications as preferred interventions given their efficacy in short-term treatment (Faraone et al., 2015).

Although evidence implies that pharmacological treatments of ADHD improve long-term outcomes but generally not to normal levels, a recent qualitative review of studies that investigated the effects of ADHD medication demonstrated an estimated 9% to 58% decline in adverse outcomes, such as injuries, motor vehicle accidents, inadequate academic achievements, and substance use disorder (Chang et al., 2019). Relatedly, a recent systematic review and meta-analysis assessing the effects of ADHD medication on associated functional outcomes suggested that there is strong support for the protective effects of stimulants against costly and highly morbid adverse functional outcomes in individuals with ADHD, mainly when individuals showed compliance with their medication treatment (Boland et al., 2020).

Nonpharmacological interventions for ADHD, otherwise known as psychosocial interventions, are part of the multimodal approach. Although psychosocial treatments usually do not decrease the core symptoms of ADHD, they are essential to improve children's functioning across multiple settings (Schoenfelder & Sasser, 2016). There are well-established psychosocial interventions targeting ADHD (Schatz et al., 2020). Many of these interventions focus on behavior modification procedures that involve engaging caregivers, parents, and teachers to apply behavioral contingencies into the child's environment, such as home and classroom, to modify the child's behavior by increasing the desired behavior and diminishing undesirable behavior (Fabiano et al., 2009; Faraone et al., 2015).

Recognizing the Needs of Children with Subclinical ADHD

The evidence-based medical approach is an example illustrating that decisions about treatment of ADHD are not exclusively based on a categorical approach. Collecting multiple clinical data is an essential step in the evidence-based approach to determine whether an individual has a categorical diagnosis (Jarrett, Meter, Youngstrom, Hilton, & Ollendick, 2018). Using this approach, clinicians will provide costly assessment procedures and treatments for individuals whose clinical data estimate the likelihood of a categorical disorder. On the other hand, those with a high probability of the disorder but still not reaching the clinical threshold might be offered less expensive and intense treatment with fewer adverse effects (Coghill & Sonuga-Barke, 2012).

Relatedly, the dimensional approach would help clinicians understand the heterogeneity of ADHD symptoms' severity and impairments by providing effective treatments for those at the diagnostic threshold and also recognizing the needs of children with subclinical disorders and providing them with clinical recommendations, appropriate referrals, and future monitoring

(Frick & Nigg, 2012). Nonetheless, it is worth noting that this recognition might have some risks. For instance, where to place a cutoff to separate children who should be considered a subthreshold from those who are not might create confusion and lead to assigning multiple subjective, inconsistent, and random thresholds (Frick & Nigg, 2012). In addition, when many individuals are included in a subclinical group, they might experience the stigma of having a psychiatric disorder, yet they may not be eligible to receive reimbursement for interventions by health-care-insurance carriers (Frick & Nigg, 2012).

Diagnostic Procedures for ADHD and Dimensional Rating Scales

An accurate diagnosis is the first step toward effective treatment. A systematic and standardized method to assess ADHD is necessary to make a valid diagnosis and prevent under or over diagnosis of the disorder. Clinical guidelines for ADHD evaluation offer systematic, thorough, and descriptive evaluation practices for clinicians. In the United States, current clinical guidelines are published by the American Academy of Child and Adolescent Psychiatry (AACAP), the American Academy of Pediatrics (AAP), and most recently by the Society for Developmental and Behavioral Pediatrics (SDBP), which emphasizes assessment and treatment of complex ADHD. All these guidelines recommend using dimensional assessment tools to support an ADHD categorical diagnosis. The following sections give a brief summary about the ADHD diagnostic process, and the advantages and limitations of including dimensional rating scales in this process.

Complexity of ADHD Diagnostic Procedures

ADHD diagnosis is a complex process with numerous challenges (Parker & Corkum, 2016). When evaluating and treating children with ADHD, clinicians must be prepared to consider a wide variety of disorders that may overlap with ADHD and coexisting conditions.

ADHD is often associated with other disorders, and behavioral comorbidity is the rule rather than the exception (Parker & Corkum, 2016). The most common ADHD comorbidities are oppositional defiant disorder, conduct disorder, learning disorders, anxiety and affective disorders, autism-spectrum disorders, and tic disorders (Pham & Riviere, 2015; Pliszka, 2003). The prevalence of these disorders along with ADHD is even greater in clinically referred children with ADHD (67%–87%) (Reale et al., 2017). The clinical picture of children with ADHD is complicated by the fact that there are many medical and psychiatric disorders that are similar to or that “mimic” symptoms that are manifest in primary ADHD; these include epilepsy, thyroid disorders, sleep disorders, drug interactions, and anemia, among others. To put it concisely, though it seems simple to “rule in” ADHD symptoms, it is challenging to “rule out” other causes of these symptoms (Parker & Corkum, 2016).

Children with ADHD Symptoms and Challenging Clinical Profiles

The AAP recommends that children with ADHD symptoms and challenging clinical profiles of possible comorbid conditions be referred to clinicians with experience in specialized care for ADHD if the clinician who first evaluated the child does not have the adequate training in the initial treatment of those conditions. The SDBP guidelines support this AAP recommendation and also emphasize that collaboration of multidisciplinary teams is required to provide a valid diagnosis and best-evidence-based treatment to children with complex ADHD. Acknowledging the complex clinical profile of children with symptom overlap or with comorbid conditions, the SDBP developed a guideline that focuses on diagnosis and treatments for complex ADHD. According to the guideline, complex ADHD is “defined by age (<4 years or presentation at age >12 years), presence of coexisting conditions, moderate to severe functional impairment, diagnostic uncertainty, or inadequate response to treatment” (Barbarese, 2020, p. 1).

Children with complex ADHD need advanced care by specialized clinicians because their evaluations require longer visits and treatment plans that take into consideration prevention of severe long-term outcomes.

Assessment of ADHD

For diagnosis of ADHD, most guidelines require a comprehensive evaluation that includes assessment of the *DSM* symptoms, the child's impairment across settings, family history, family functioning, the child's perinatal history, developmental milestones, medical history, comorbid disorders, physical disorders, and developmental conditions. Most guidelines suggest a diagnostic interview (unstructured or structured) with parents or caregivers to make a categorical diagnosis of ADHD. The AACAP also recommends that clinicians interview children and adolescents separately from their parents, not for the purpose of making a diagnosis but for evaluating other comorbid symptoms (e.g., anxiety or depression) that might not be revealed by the child when his or her parents are present in the room (Pliszka, 2007).

Dimensional Rating Scales

It has been debated whether there is a need for rating scales if the diagnosis of ADHD depends on clinical expertise. Although rating scales are not considered sufficient for clinical ADHD diagnosis, they are endorsed as good clinical practice for ADHD screening purposes, for measuring symptom severity, and for treatment monitoring (Collett et al., 2003; Pliszka, 2007). ADHD behavioral rating scales hold to the premise of dimensional conceptualization of ADHD. They are designed to measure a range of quantitative degrees of symptom severity. Historically, using behavioral rating scales is one of the early methods that provided dimensional adjuncts to ADHD categorical diagnoses (Swanson et al., 2009). Since the late 1960s, behavioral rating scales have been used in research studies to help describe and diagnose individuals with ADHD-

like behaviors (e.g., Goyette, Conners, & Ulrich, 1978). ADHD rating scales often involve a number of items such as closed-ended questions or statements that ask an informant (e.g., parent, teacher, or child) to rate retrospectively the child's ADHD symptom frequency, intensity, and severity. However, the use of quantitative scores does not disregard the use of categorical decisions. Rating scales provide scores on ADHD dimensions. The use of cutoff points as a basis for yes-or-no decisions can assist in ADHD clinical categorical diagnosis for differentiating between cases and non-cases (Hudziak, Achenbach, Althoff, & Pine, 2007).

Typically, there are two approaches to ADHD rating scales that have been used as dimensional supplements to complement categorical diagnosis. The first type consists of items that are not based on a symptoms list of diagnostic classifications from standard nosological systems (e.g., *DSM*). Instead, they are based on applying multivariate statistical analyses (e.g., factor analytic method) to empirical data collected from representative samples of children to determine which symptoms tend to occur together as syndromes. Thus, these scales are labeled as “from the ground up” scales (Swanson et al., 2009). Some of these scales are broadband (i.e., assessing various behavioral conditions related to general psychopathology, along with ADHD-related symptoms). The broadband behavioral rating scales are generally used for evaluating treatment outcomes because they are sensitive to both behavior and pharmacological treatment effects (Pelham, Fabiano, & Massetti, 2005). Although scales have been found to agree well with ADHD *DSM* diagnosis (Chang, Wang, & Tsai, 2016), some guidelines do not recommend using scales for diagnosing ADHD in clinical practice (AAP, 2019).

The second approach to ADHD scales is usually described as “top-down” or narrowband (i.e., ADHD-specific) because rating-scale items are developed based on the ADHD symptoms from the *DSM*. These items provide frequency or severity indicators of the *DSM* symptoms list.

Evidence suggests that using these scales improves the sensitivity and specificity of the categorical diagnosis of ADHD based on the *DSM* instead of on scales based on general psychopathology (Swanson et al., 2009). The strong association between these rating scales and the *DSM* diagnosis is not surprising since items of these rating scales are defined based on *DSM* items. Nonetheless, there are not that many differences between the broadband behavioral rating scales (e.g., Conners Scale or CBCL attention subscale) and *DSM* scales; both are moderately correlated with ADHD symptoms (Chang et al., 2016).

Advantages of ADHD Rating Scales

Most ADHD guidelines recommend applying rating scales to gather information from different informants (e.g., parent, teacher, and caregiver). Behavioral rating scales are routinely used in clinical practice to help clinicians assess a child's behavior across various settings from multiple perspectives in a cost-effective way as they are inexpensive relative to other diagnostic procedures. Evidence shows that they are useful to measure treatment outcomes and are markedly sensitive to both behavioral and pharmacological treatment effects (Collett et al., 2003; Pelham et al., 2005). The AACAP guidelines recommend a pre-diagnosis screening by incorporating direct questions or providing rating scales covering the *DSM* symptoms and impairments related to these symptoms during any mental health assessment. These rating scales could be completed prior to the visit to indicate whether a comprehensive evaluation is needed if the child scores in the clinical range for ADHD diagnosis, and if impairments related to ADHD symptoms have been reported. AACAP explicitly provides a list of recommended scales and encourages the use of rating scales that measure both *DSM* symptoms and other possible comorbid disorders. AAP supports the use of rating scales based on *DSM-5*.

Because psychometric principles are the basis for any behavioral rating scales, their data about severity level and impairments of children with the disorder can be quantified and analyzed (Verhulst & Van der Ende, 2006). Thus, in comparison to categorical methods, behavioral rating scales provide more statistical information. An additional advantage of behavioral rating scales is that the majority of these scales apply standardized methods for selecting age- and gender-sensitive cutoffs (Collett et al., 2003; Swanson et al., 2009). Data for these scales are collected both for children with ADHD-related symptoms and for randomly selected peers of the same age and gender to provide normative quantitative indicators regarding severity of a child's ADHD symptoms compared to group means (Collett et al., 2003; Pelham et al., 2005; Swanson et al., 2009).

Limitations of ADHD Rating Scales

A major limitation of the most common ADHD rating scales is that they are susceptible to "rater bias," such as when one rater's report consistently tends to over- or underestimate ADHD symptoms. Thus, there is a need for obtaining ADHD rating reports from multiple informants (e.g., parents and teachers and self-report). Nonetheless, multiple informants do not always agree, and there is a risk of lack of concordance between informants due to rater biases and accuracy factors.

Although behavioral rating scales probably enhance the precision of the diagnostic process when normative data and cutoff scores are available to help clinicians determine whether a child's symptoms score is "statistically deviant or not," it must be pointed out that not every statistical deviance indicates clinical significance (Shekim et al., 1986, p. 657). Concerns have been raised regarding the distribution of scores of most behavioral rating scales when data are collected from the general population, namely, their tendency to yield skewed rather than normal

distributions (Swanson et al., 2012; Swanson et al., 2009). The reason for that is that, when applying statistical cutoffs based on population norms when the normal distribution assumptions are not met, “psychometric flaws” likely will present.

Most ADHD rating scales quantify ADHD symptoms on a 4-point (0 to 3) Likert scale to capture the degree of presence, intensity, and frequency of a problem, with supplementing verbal anchors (e.g., 0 = Never, Rarely, 1 = Sometimes, 2 = Often, or 3 = Very Often). The problem with this strict range of scores is assigning a zero value to all children who have none of the ADHD symptoms. It also gathers them in one group as “normal” children without psychopathological problems and ignores variations in their attention and behavioral control abilities. Thus, these scales often create an exponential rather than a normal distribution when applied to population-based or epidemiologic samples. A skewness in scores can occur because significant problems in attention and behavioral control abilities are rare in the population, and most children rated between having few symptoms (Sometimes = 1) or none at all (Never = 0). Thus, most of the summary score results of these rating scales are likely to take a *J* shape distribution due to discontinuity of measuring variations in ADHD dimensions beyond the diagnostic threshold in population samples. In other words, none of these scales capture the entire distribution of ADHD dimensions, which is in opposition to the assumption of a normal distribution of attention and behavioral control abilities in the population. Despite that, most common rating scales use linear transformations of the raw scores to set cutoff points derived from population norms (e.g., two standard deviations above mean + 2 *SD* ($z = 2$), *T*-score = 70, 95th percentile) to define moderate to severe cases of ADHD that require clinical attention on the extreme end of ADHD dimensions. Consequently, these rating scales share the same risk

because when the normal distribution assumption is violated, it is problematic to apply such statistical cutoffs (Swanson et al., 2012).

An additional problematic consequence of truncated distributions demonstrated by most common ADHD rating scales in the general population is reducing the statistical power needed to identify ADHD genetic markers (Brites, Salgado-Azoni, Ferreira, Lima, & Ciasca, 2015 ; Burton et al., 2019; Greven et al., 2018). For instance, molecular genetic studies assume that ADHD is a highly polygenic disorder which is influenced mainly by additive common gene variants, each having a small effect (Thapar, 2018). Specifically, this approach assumes that cases with ADHD and controls differ quantitatively because the genetic component to ADHD “disease-associated alleles” may also induce considerable phenotypic variation among the general population. However, as mentioned earlier, most ADHD rating scales offer limited information about ADHD behavioral variations distributed among the general population who expected to report rare extreme ADHD symptoms and, therefore, those scales have diminished power to detect genetic effects because they do not allow for comparisons between extreme groups (Burton et al., 2019; Coghill & Sonuga-Barke, 2012; Faraone & Larsson, 2019; Greven et al., 2018).

A further drawback of the most common ADHD rating scales is that they are grounded in a “problem-based” approach. They ask informants to report what is wrong with a child’s attention and behavioral control abilities based on deficit-oriented items, and they disregard recording the child’s strengths on such abilities, if any. The “problem-based” approach is valuable from a diagnostic sense because it helps identify moderate to severe cases of children with ADHD who need intervention. However, these scales can be described as “unidirectional” and “symptom severity” rather than “full-continuum” behavioral rating scales. They only

measure severe to moderate symptoms at one end of the ADHD distribution and ignore existing differences among less symptomatic and normally functioning individuals in the general population.

Part II:

The Strengths and Weaknesses of the SWAN Scale

With the goal to overcome limitations of the current ADHD rating scales, Swanson et al. (2012) developed the Strengths and Weaknesses of Attention-Deficit/Hyperactivity Symptoms and Normal Behaviors Scale (SWAN). The SWAN is a narrowband rating scale that contains 18 items that correspond to the ADHD diagnostic criteria of the *DSM-5*, and it has two subscales representing the two-factor structure of ADHD. The SWAN has multiple characteristics that make it different from most other ADHD rating scales (Brites et al., 2015). Detailed descriptions of these characteristics are provided in the following sections.

Integrating the Continuum Approach to ADHD and The DSM Top-Down Approach

Although most ADHD rating scales are based on the dimensional approach to ADHD, they focus only on measuring symptom severity and do not measure the variations on ADHD dimensions beyond the cutoff of absence of the disorder. Swanson et al. (2012) designed the SWAN based on a “continuum of behavior” approach to ADHD. The SWAN was designed to capture scores ranging across an entire theoretical continuum of ADHD dimensions that might be distributed normally in the population. The SWAN’s goal is to identify individuals with strengths (e.g., above-average attention and activity control skills) and those with weaknesses (e.g., inattention and hyperactivity–impulsivity that indicate a clinical ADHD diagnosis) along ADHD behavioral dimensions. Although the SWAN was created to offer a tool to assess ADHD based on a continuum approach, the *DSM* top-down approach was also considered to ensure

agreements between the SWAN's scores and ADHD clinical diagnosis based on *DSM* criteria. In order to bring these two approaches together, Swanson et al. (2012) used the content of the Swanson, Nolan, and Pelham's (SNAP-IV) rating scale—a narrowband, 4-point rating scale, which includes 18 items based on the *DSM-IV* ADHD symptoms.

However, Swanson et al. (2001) applied some modifications to the SNAP-IV rating scale so that the SWAN new items could capture variations in the entire continuum of ADHD dimensions. Items of the SNAP-IV were converted from negative statements (e.g., does not seem to listen when spoken to directly) to more neutral statements (e.g., listens when spoken to directly). As a result of changing each item of the SNAP-IV to a bidirectional statement that allows reporting of strengths and weaknesses, each item of the SWAN can capture variations along the entire underlying ADHD dimensions of behavior at every item level. Further, instead of using the 4-point scale of the SNAP-IV, which measures only the pathological severity of ADHD, the SWAN was based on a broader range of scores to rate a child's behavior on a 7-point scale. Each of the SWAN items rates a child's behavior from -3 to $+3$ (below average to above average), and 0 is set up as “normal” according to the population average.

Theoretical Statistical Cutoff

The refinements of rephrasing the SNAP-IV items and using the 7-point scale were employed to yield a distribution that approximates the normal distribution in the population, which should reflect the entire underlying range of ADHD behaviors. Given that SWAN score distributions are designed to meet the assumption of normality, the SWAN uses theoretical z -scores to estimate cases of ADHD in the population. Conversely, most ADHD rating scales make adjustments to their z -scores or T -scores or use empirically derived, rather than theoretical,

percentiles to interpret their scores and to estimate cases of ADHD because their score distributions in the population are usually truncated (not normally distributed).

Swanson et al. (2012) pointed out that the high prevalence of cases of ADHD in the population could be one detrimental byproduct of using a 4-point scale of ADHD severity and statistical cutoffs from a non-normal (skewed) distribution. To examine this issue, they collected teacher ratings of ADHD symptoms for children between kindergarten to grade six ($N = 847$) based on the SNAP-IV and the SWAN. The distribution of the SNAP-IV summary scores for their sample exhibited extreme rightward skewness, as most of the children (79.9%) scored one or less (“Just a Little” or “Not at All”) on the ADHD symptoms. They used a theoretical statistical cutoff ($z = 1.65$) in examining the effect of applying cutoffs that assumed a normal distribution when the normal distribution assumption is not met. They questioned whether that technique might be inappropriate and lead to higher prevalence of ADHD than expected. This theoretical cutoff, if the normal distribution assumption is met, is supposed to classify 5% (the average rating of ADHD in school-aged population of children) of the sample. But instead, this cutoff, based on the SNAP-IV scores, classified more cases than would be expected, by almost 1.7 times (8.4%). On the contrary, when they used the z -score cutoff of 1.65 SD to calculate the ADHD prevalence based on the SWAN scores, they identified less than the expected 5% estimation of extreme cases with ADHD.

Strength-Based Assessment Approach of ADHD

An additional advantage of using neutral language for the SWAN items is implementing a *strength-based assessment* approach, which might minimize the stigma associated with reporting only a child’s problems by allowing the informants to report the child’s strengths, if any (Polderman et al., 2007; Robaey, Amre, Schachar, & Simard, 2007). This approach has

already been implemented by the Strengths and Difficulties Questionnaire (Goodman, 1997). The SDQ has 25 items that assess various behavior and emotional problems associated with many psychopathological disorders involving ADHD. Some of the SDQ items assess strengths (e.g., sees tasks through to the end, and good attention), and others assess problems (e.g., constantly fidgeting or squirming). Thus, although both the SDQ and the SWAN assess a child's strengths, the SDQ allows the informants to report child strengths only in some aspects of a child's behavior. On the other hand, the SWAN uses bidirectional statements that allow rating strengths and problems simultaneously for all of *DSM-5* symptom criteria.

The SWAN as a Research Instrument for ADHD Traits

The unique characteristics of the SWAN make it an appealing tool for researchers interested in measuring variations along the full ADHD continuum in the general population. Several studies compared the SWAN scores distribution in the population to other ADHD rating scales, supporting the claim of the SWAN to produce an approximately normal distribution in both clinical (Chan et al., 2014; Hung, & Leung, 2014) and population (Burton et al., 2019; Greven et al., 2018; Lai et al., 2011) samples. Consequently, the SWAN has the advantage of increasing the statistical power needed for research studies because it measures variations of ADHD dimensions even among normally functioning individuals. The SWAN has been used in various genetic studies (Burton et al., 2019; Hay et al., 2007; Sergeant, & Swanson, 2007; Polderman et al., 2007). It has also been used to examine some neuropsychological aspects of ADHD and to assess response to medical stimulants and other trials for ADHD interventions, such as efficacy of neurofeedback (Wang et al., 2021), family-based mindfulness (Siebelink et al., 2021), and elimination diets (Bosch et al., 2020).

Psychometric Properties of the SWAN

Many studies examined different indexes of psychometric properties of the original English version of the SWAN and the translated forms of it in many languages such as Dutch, French, Spanish, and Chinese (see Table 2). These studies reported excellent psychometric properties of the SWAN, such as that its scores were normally distributed in population samples (Arnett et al., 2011; Greven et al., 2018; Schulz-Zhecheva et al., 2017). Many studies also demonstrated that the SWAN has high internal consistency and Cronbach's alpha coefficients range between $\alpha = 0.87$ and $\alpha = 0.95$ (Arnett et al., 2011; Lakes et al., 2011; Schulz-Zhecheva et al., 2017), moderate retest reliability (Lai et al., 2011; Lakes et al., 2011), and stable two-factorial structure (Arnett et al., 2011; Gomez, Vance, & Gomez, 2013; Lai et al., 2011). Detailed description about the SWAN's diagnostic validity and ranges of specificity and sensitivity values will be addressed in the following sections.

The SWAN Clinical Utility

Although a rating scale must be supported by strong psychometric evidence, to be incorporated into the assessment process it must also be clinically useful. There is limited support for the SWAN's utility in clinical practice as a measure for identifying cases of ADHD (Brites et al., 2015; Chan et al., 2014; Collett et al., 2003 2003; Robaey et al., 2007). Clinical utility of an assessment measure has been defined in many ways. However, the key indicators of the usefulness of a rating scale are its diagnostic accuracy and its ability to provide information that improves everyday clinical decision-making (Hunsley, 2004).

The following sections include further details regarding studies that investigated three psychometric indicators of the SWAN that are crucial to determining whether the SWAN has clinical utility similar to that of other validated rating scales that are commonly used in clinical

practice. First, I will review studies that investigated the convergent validity of the SWAN with other ADHD rating scales. Next, I will describe the literature related directly to the degree of correlation between the SWAN scores and a clinically relevant construct—functional impairments. Last, I will examine studies that investigate the diagnostic validity of the SWAN.

The Convergent Validity of the SWAN with ADHD Symptoms Measures

Convergent validity refers to the degree of correlation between two scales that measure similar constructs. It implies that two scales that measure the same construct should be highly correlated (Carlson & Herdman, 2010). Although convergent validity is a common indicator of construct validity, it does not convey that converged scales are measuring the construct they are intended to measure. Convergent validity indicates only that converged scales collect similar information; thus, they are expected to produce comparable research results (Carlson & Herdman, 2010). Evidence of convergent validity is established based on the magnitude of the correlation between a scale of interest (e.g., the SWAN) and other scales (not necessarily a gold standard) that measure conceptually related constructs (Echevarría-Guanilo, Gonçalves, & Romanoski, 2019). Although a correlation coefficient above 0.5 is adequate to establish convergent validity, a score above 0.70 is usually recommended (Carlson & Herdman, 2010).

In terms of the SWAN convergent validity with narrowband scales that focus on the *DSM*-oriented ADHD symptoms, a recent study by Burton et al. (2019) revealed that the SWAN scores showed high convergent validity with the Conners scale as evidenced by a correlation coefficient of 0.72 in a community sample of children aged 6–17 years ($N = 5,154$) (Burton et al., 2019). Also, results from a study by Robaey et al., (2007) in a sample of clinic-referred children ($N = 88$) aged 6–9 years revealed highly significant correlations (-0.79, -0.85) between the French SWAN subscale scores and the Conners inattention and hyperactivity–impulsivity

subscale scores (Robaey et al., 2007). The SWAN also showed low to moderate correlations (0.48 to 0.53) with scores on the Disruptive Behavior Rating Scale (DBRS) in a large sample ($N = 1,502$) of twins aged 5–7 years in the United States and Australia (Arnett et al., 2011).

In terms of the SWAN's convergent validity with broadband scales, four studies estimated the SWAN's convergent validity with two subscales of broadband scales: Child Behavior Checklist–Attention Problem (CBCL-AP) and SDQ/HI (Table 2). These subscales are helpful during the screening phase because they typically have fewer items and are part of measures that assess various behavior and emotional problems associated with many psychopathological disorders involving ADHD. Overall findings from these four studies, all based on community samples, revealed low to moderate correlations (0.40–0.53) (Lakatos, Birkás, Tóth, & Gervai, 2010a; Polderman et al., 2007; Schulz-Zhecheva et al., 2017) between the SWAN and CBCL-AP and low to high correlations (0.49–0.74) between the SWAN and SDQ/HI. Findings from a study by Lakatos et al. (2010) showed that the Hungarian SWAN subscales had higher correlations (0.67–0.74) with the Hungarian SDQ/HI than with CBCL-AP (0.40–0.49) based on a sample of ($N = 156$) children aged 6 years.

The Concurrent Validity of the SWAN with Functional Impairment Measures

Swanson et al. (2011) emphasized the need for concurrent assessment of impairment when assessing ADHD through a continuum approach that does not replace the clinical judgment about the need for treatment. The SWAN scores' association with functional impairment scales is still understudied. One study examined this association in a community sample of 280 Iranian children aged 6–12 years; a group analysis was conducted to compare functional impairment of children with strong, moderate, and weak attention and behavioral control based on SWAN scores. Results of the study revealed that parents who reported on the SWAN that their children

had more ADHD symptoms also rated their children higher on the Weiss Functional Impairment Rating Scale–Parent Report (WFIRS–P) (Kiani, Hadianfard, & Mitchell, 2019 2019). However, the WFIRS is less likely to be used in clinical practice given that it is lengthy (70 items) and expected to be time consuming for completion in a clinical setting, which might place a burden on raters. So far, no study has examined the extent to which the SWAN correlates with a brief functional-impairment scale, in a clinical sample, to provide further evidence for the SWAN’s usefulness as a clinical diagnostic tool.

The Diagnostic Validity of the SWAN

Diagnostic accuracy refers to the degree of agreement between a rating scale (e.g., the SWAN) with a reference or gold standard (Cohen et al., 2016). Two common methods to quantify the diagnostic accuracy of a scale are sensitivity (the ability of a scale to correctly identify ADHD cases as positive) and specificity (the probability of a scale’s correctly identifying non-cases as negative). Receiver-operating-characteristic (ROC) curve analysis is useful in estimating the extent to which a screening measure can identify a clinical disorder (such as ADHD) that is defined dichotomously as present or absent using an established reference standard. Generally, the ROC analysis is helpful in identifying cutoffs corresponding to balanced levels of scale sensitivity and specificity (Youngstrom, 2014). Also, the ROC analysis allows researchers to calculate sensitivity and specificity for all possible cutoff points of a score and then combine them in a single value called “area under the curve” (AUC) (Youngstrom, 2014). The AUC may have any value between 0 and 1: a value of 1 indicates the scale’s ability to perfectly discriminate, and a value of 0.5 (Youngstrom, 2014) or less reflects the inability to discriminate (Rice & Harris, 2005).

A few published studies have investigated the diagnostic validity of the SWAN (see Table 2). Three studies investigating the SWAN total and subscale diagnostic accuracy used community samples (Burton et al., 2019; Lai et al., 2011; Schulz-Zhecheva et al., 2017), and their results were based on distinguishing children with ADHD from healthy controls (i.e., without other psychiatric disorders).

Schulz-Zhecheva and colleagues (2017), examined the diagnostic utility of the German version of the SWAN subscales and combined score in differentiating a clinic ADHD sample ($n=62$) from a community sample ($n=343$) of school-age children aged between 8 and 18 years. The ROC analysis was used to determine the ability of SWAN to discriminate between ADHD cases in the clinical sample and healthy participants without any present neurological or psychiatric disorders in the community sample. Results of the ROC analysis indicated that the SWAN has an excellent ability to differentiate between children with and without ADHD. Specifically, the SWAN total score had very good discriminant ability: AUC value (0.91), sensitivity (0.84) and specificity (0.88).

Lai and colleagues (2011) examined the ability of the Chinese version of the SWAN subscales and combined score in differentiating a clinic ADHD sample ($n=247$) from a community sample ($n=3,722$) of children aged between 6 and 12 years in Hong Kong. Results from the ROC showed that the SWAN subscale and combined scores all had good discriminant validity in differentiating the ADHD clinic sample from the community sample. Lai and colleagues (2011) also highlighted an important point related to the SWAN discriminant ability based on the cutoff scores at 1.65 *SD* above the mean as suggested by Swanson et al. (2005). When they applied a cutoff score (1.65 *SD*) to their community sample, although that resulted in high specificity of the SWAN scale (0.95), the sensitivity of the SWAN was lower (0.42).

Because both specificity and sensitivity are equally important, it is not recommended to compromise one in favor of the other in ADHD screening. Thus, they calculated different cutoff scores by balancing the sensitivity and specificity, using the sum of both. These cutoffs resulted in sensitivity that ranged between 0.67 and 0.83 and specificity that ranged between 0.76 and 0.89.

Burton and colleagues (Burton et al., 2019) collected parent-report SWAN scores for a large community sample that included 12,797 children aged 6–17 years and established cut-points for the SWAN based on age and sex. Diagnosis of ADHD was reported for 6.2% of this community sample based on the optimal cut-point (>0.74) identified in the ROC analyses for SWAN-Total score. Also, the ROC analysis showed high sensitivity and specificity of the SWAN (AUC = 0.85–0.88). Further, Burton and colleagues (2019) validated the SWAN cut-point of (0.74) based on their community sample to distinguish cases with ADHD diagnosis in a clinical sample of 266 ADHD cases group and a control group of 36 healthy children- both groups were children aged 6–17 years. In the clinical sample, Burton et al. (2019) reported that the cut-point of 0.74 identified 222 ADHD cases out of 266 and indicated that the SWAN had high sensitivity (0.84) and specificity (0.92). However, when the cutoff 1.65 SD was applied, the specificity of the SWAN reached 100%, while the sensitivity dropped significantly to 0.36. Similar to the findings of Lai and colleagues (2011), both results suggest that Swanson's (2012) suggested cutoff point of (1.65 SD) compromised the sensitivity over the specificity of the SWAN score to discriminate between ADHD cases and healthy controls.

Although these studies provide evidence about the SWAN's ability to discriminate ADHD cases from children in community samples, they provide little evidence for the SWAN's clinical utility. The SWAN scores were compared between a sample of confirmed ADHD cases

and healthy controls in each of these three studies. Typically, children referred for ADHD evaluation are at risk for high levels of impairments and comorbidities compared to their community counterparts. Thus, using highly restricted samples of healthy controls yields less clinically applicable findings because in their practice, clinicians are most likely to have to decide whether a child has ADHD, some other diagnosis, or comorbid diagnoses (Raiker et al., 2017). As yet, only two studies (French and Chinese translated forms), described next, examined the SWAN's diagnostic validity to differentiate ADHD from other psychiatric diagnoses (non-ADHD) among clinical samples of children.

In the study by Robaey et al. (2007), in a sample of clinic-referred children aged 6–9 years, the ROC curves were generated to compare the SWAN total diagnostic accuracy to the Conners Rating Scale using a structured interview, the Diagnostic Interview Schedule for Children Version IV (DISC-IV), as reference for the presence or absence of ADHD diagnosis. The SWAN-French total score showed a better sensitivity (0.86) and specificity (0.88) than the Conners overall score sensitivity (0.75) and specificity (0.80). Also, the SWAN-French had a larger AUC (0.89) than the AUC of the Conners AUC (0.79). However, these results were limited by the small sample ($N = 88$).

The second study, by Chanet et al. (2014), compared the Chinese-SWAN with the ADHD diagnosis obtained from the DISC-IV. ROC analyses revealed that the Chinese-SWAN subscales (SWAN-IA, SWAN-HI) effectively differentiated ADHD from non-ADHD psychiatric controls (AUCs of 0.94 and 0.95, respectively) among 290 children aged 6–12 years in a psychiatric clinic setting. However, no information was reported in this study about the SWAN total score cutoff point, AUC, sensitivity, or specificity (Chan et al., 2014).

Overall, these two studies provided preliminary evidence of the SWAN total and subscales' capacity to discriminate between ADHD and other disorders in clinical settings. However, both studies used parent-structured interviews (DISC-IV) as references for ADHD diagnosis, which might increase the risk of shared method variance between the SWAN scores and clinical diagnosis, as both were obtained from parents. Therefore, results from these studies might also overestimate the SWAN's usefulness in clinical settings. To better estimate the SWAN total and subscales' diagnostic accuracy in a clinical population, a reference standard based on a best-practices comprehensive assessment that integrates information from multiple sources (e.g., semi-structured interviews with the parent and the child, developmental history, and teacher rating scales) should be used. As of yet, no study has evaluated the SWAN-English's diagnostic validity or established cutoff points for the SWAN total and subscales for a clinic-referred sample based on the current *DSM-5* criteria and best-practice comprehensive assessment for ADHD diagnosis.

Summary

Over the past decades, the conceptualization of ADHD has continued to be refined. Current practice is to define ADHD categorially and dimensionally as the extreme end of a continuum and that individuals with ADHD differ from those without ADHD by degree rather than in kind. Researchers in the field are moving toward an integrated approach that incorporates dimensional approaches of ADHD over time to support ADHD categorical decisions and to increase its applicability to clinical decisions. The SWAN is an example of ADHD behavior rating scales that are based on the *DSM* criteria and are designed to measure the full continuum of ADHD behavior including potential positive aspects. Nonetheless, little is known about the SWAN's clinical utility. Further work is needed to understand the degree to which the SWAN

corresponds to other measures of ADHD including clinical ADHD diagnosis and to functional-impairments ratings. This study addresses these gaps by evaluating the SWAN's clinical utility compared to a well-established commonly used measure (SDQ/HI).

Chapter 3: Methods

The overarching goal of this study was to examine the clinical utility of the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) scale, defined by its ability to classify children referred to a child psychiatry outpatient clinic into ADHD and non-ADHD cases. This chapter presents an overview of the research design and procedures for the current study. Also included are descriptions of the study data preparation, participants, operational definition of the ADHD criterion diagnosis, study measures, and the statistical analysis plan.

Study Design

This study used a cross-sectional cohort design to assess and compare the performance of ADHD rating scales: the SWAN and SDQ/HI for identifying cases of ADHD. It utilized archived data of children who visited a children's hospital's mental health outpatient clinic, specializing in ADHD and related disorders, in the northwest United States. All study participants were referred by community pediatricians, schools, and mental health professionals between January 1, 2015, and August 31, 2019. All the study participants were referred to the clinic based on a suspicion of their having attentional and behavioral problems; therefore, all participants received a comprehensive clinical evaluation to determine their diagnoses. The three parent-reported rating scales included in this study were part of a pre-visit online screening as part of the clinical assessment routine. One parent/caregiver completed the three rating scales, either on the same day or within a short time period, on average 3 days before their child's first visit to the clinic for a comprehensive evaluation for ADHD.

Data Preparation

The study was approved by the institutional review board (IRB) at the study site. At the same time, an IRB approval from the Human Subjects Division of the University of Washington

was obtained for this study under the cooperative agreement between the study site and the University of Washington/Seattle. This study posed only a minimal risk to parents and children because it was a chart review of retrospective data, and there were no physical procedures involved. Although there was a risk of breach of confidentiality, multiple steps were applied to secure the data and minimize the risk of inadvertent disclosure.

Based on the study inclusion/exclusion criteria (see Figure 3.1), data for the current study were extracted from two sources: the online database for the pre-visit screening routine of the clinic and participants' medical records. Data from the study's three parent-reported measures were obtained from a web-based online library that included a screening battery with multiple rating scales assessing caregiver perceptions of children's different behavioral and emotional problems, ADHD symptoms, and impairments. Then demographic data and clinical diagnoses for children who fulfilled the study inclusion criteria were extracted from each child's medical records. Primary demographic variables used in this study were sex, age, and race/ethnicity. The current study data files contained integrated information from both sources stored in Excel files in the study site OneDrive. This Microsoft online service provides cloud file storage resources that allow the clinic affiliates to access files. Finally, to avoid any potential of violating HIPAA, all direct patient identifiers (e.g., patients' names, medical record numbers) were removed, and data were made available to be transferred to statistical software for data analysis.

Participants

This study utilized archived data for children who visited a mental health outpatient clinic specializing in ADHD and related disorders at a children's hospital, in the northwest United States. Inclusion criteria for selecting participants' data for this study were as follows:

1. Participants were children aged 6–11 years, newly referred to the study site between January 1, 2015, and August 31, 2019. This time range was chosen because this study measures (SWAN, SDQ/HI and IRS) were collected concurrently between the year of 2015 to 2019 as part of pre-visit screening routine in the study site.
2. Only one parent/caregiver of each study participant completed all study measures concurrently, and these measures were collected online as part of the clinical routine before the child's evaluation visit.
3. All participants of the study received a comprehensive ADHD evaluation.

Exclusion criteria were minimized because this study's aim was to produce clinically applicable findings, and clinicians usually work with clinically diverse populations. To test whether any rating scale is clinically useful, the context during data collection should simulate how clinicians might use a scale in real clinical situations (Youngstrom, 2014).

Based on inclusion/exclusion criteria (see Figure 3.1), this study included data from 357 participants. Out of the 357 children, based on the diagnoses documented in the medical records of each child, a total of 257 were ADHD cases (73%), and 100 were non-ADHD (27%) (children with other psychiatric or medical diagnoses; see Table 3.1). Most of the sample were male subjects ($n = 262$, 73%), and slightly over half (53%) of the male subjects were ADHD cases. Female subjects represented somewhat more than a quarter of this sample ($n = 95$), and only 19% of them were ADHD cases. Regarding the race of the subjects, the majority were White (65%); Asians represented 11%; and Black, or African American, children represented 5% of the sample. No differences between the ADHD group and the non-ADHD group for age, sex, race, and ethnicity were found as indicated by t , Chi-Square, and Fisher's exact tests. For age, Levene's

test for equality of variances was conducted and was not significant ($p > .05$), suggesting that that the two groups were comparable.

Clinical diagnoses were coded in the medical records based on the International Statistical Classification of Diseases and Related Health Problems (ICD-10). Thus, the ADHD group consisted of children assigned the following diagnoses: ADHD combined type, ADHD predominantly inattentive type, ADHD predominantly hyperactive type, and ADHD unspecified type. The ADHD combined type was the most common diagnosis in the ADHD group (32%, $n = 115$), with the mean (SD) for age in this group, for both males and females, only differing slightly: 7.76 (1.62) years and 7.51 (1.37) years, respectively. ADHD unspecified type was the second most common diagnosis in this group (28%, $n = 100$); the mean age for males in this group was 8.25 (1.64) years, and the mean age for females was 8.35 (1.84) years. Few participants were diagnosed as the ADHD predominantly inattentive type (10%, $n = 35$), with the mean age for this group higher for females (a mean of 10.00 (1.63) years) than for males (a mean of 8.52 (1.71) years). The ADHD predominantly hyperactive type was the least common in this sample (2%, $n = 7$), and all were male subjects with a mean age of 6.71 (.76) years, which is the youngest mean age among all the groups. Less than a third of the sample were in the non-ADHD group (males, $n = 71$; females, $n = 29$), and the mean age in this group was similar for both males and females, with a mean of 8.23 (1.73) years and 8.03 (1.56) years, respectively. The non-ADHD group had a wide array of medical and psychiatric disorders. The most common diagnoses within this group were oppositional defiant disorder (8%), followed by conduct disorder (6%) and then anxiety disorder (3%). The non-ADHD group also included cases of adjustment disorder, developmental disorder of speech and language, obsessive-compulsive

disorder, social pragmatic communication disorder, intellectual disabilities, major depressive disorder, dysthymia, and Williams syndrome (a rare genetic condition).

Measures

Definition of the ADHD Criterion Diagnosis

When examining the diagnostic validity of an ADHD scale (e.g., SWAN) and its ability to identify individuals with and without the disorder, the outcomes of this scale should be compared with outcomes of a “gold standard” criterion diagnosis in the same sample. In this current study, the criterion diagnosis for ADHD was defined based on the comprehensive clinical evaluation of ADHD that follows the American Academy of Child and Adolescent Psychiatry (AACAP) guidelines for best practice (for detailed guidelines, please see Pliszka, 2007). The ADHD clinical diagnosis in this study was formed based on comprehensive assessment and diagnostic procedures consisting of the following:

1. A detailed, semi-structured interview with the child’s parent/caregiver about the DSM 5’s list of 18 ADHD symptoms, including each symptom duration, severity, frequency, at what age these symptoms manifested, and their impact on the child’s functioning; an assessment of the child’s other psychiatric symptoms, including physical or psychological trauma, and family history; and an in-depth assessment of the child’s perinatal history, development milestones, and history of medical, neurodevelopmental, and mental health conditions
2. A separate clinical interview with the child that was not intended to confirm the ADHD disorder but to examine the child’s mental status, anxiety, and depressive symptoms, as well as other psychiatric symptoms, and their impact on the child’s functioning

3. When needed (as part of the post-interview clinical evaluation), standardized behavior rating scales for ADHD and other disorders, filled out by parents and the child
4. Teacher rating scales mailed to each child's teachers following the parents' semi-structured interview

As per standard practice in the study site clinic, diagnostic decisions were made after discussion with a multidisciplinary team (e.g., clinical psychologists, child psychiatrists) specializing in ADHD and related disorders. This comprehensive assessment was completed for each patient following two visits to the clinic (a 90-minute first visit and a 60-minute second visit) by licensed Ph.D.-level clinical psychologists or child psychiatrists. Thus, it was assumed that the ADHD diagnosis in this study obtained from each child's medical record was formed when a clinician gathered and integrated all data needed to inform the clinical diagnosis. In addition, children in the ADHD group were defined as such if they met ADHD criteria regardless of having other comorbid disorders. As put forth by the AACAP guidelines, "When patients with ADHD meet full DSM-IV criteria for a second disorder, the clinician should generally assume the patient has two or more disorders and develop a treatment plan to address each comorbid disorder in addition to the ADHD" (Pliszka, 2007, p. 901).

Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale

The SWAN scale (Swanson et al., 2012) is a dimensional ADHD rating scale that was designed to measure the full range of behavior rather than only the pathological symptoms of ADHD. The SWAN items are based on the DSM-5's ADHD symptoms criteria. However, each criterion was converted from negative statements (e.g., "is forgetful in daily activities") to more neutral statements (e.g., "remembers daily activities") to measure variations in attention skills and behavior control abilities of the general population.

In the SWAN scale, parents are asked to quantify each item regarding the severity of their child's ADHD symptoms compared with other children of the same age over the past 6 months. The SWAN is a 7-point Likert-type scale ranging from 0 (far below average) to 6 (far above average). The SWAN has 18 items, with items 1–9 corresponding to the attention deficit subscale and items 10–18 representing the hyperactivity-impulsivity subscale. A child's total score on the inattention and hyperactivity-impulsivity subscales of the SWAN scale is then averaged. Each item is scored from –3 to +3 (below average to above average), where 0 is “Average” and based on the population average. This means that positive scores imply a parental report of below-average attention skills and behavior control abilities or ADHD symptoms, whereas negative scores imply above-average attention skills and behavior control abilities.

The SWAN has been translated into many languages, including French (Robaey, Amre, Schachar, & Simard, 2007); Chinese (Lai et al., 2011), Spanish (Lakes, Swanson, & Riggs, 2011), and Dutch (Schulz-Zhecheva et al., 2017). Evaluation studies of SWAN psychometric properties across community and clinical samples have reported that it has excellent internal consistency, and Cronbach's alpha coefficients range between $\alpha = 0.87$ and $\alpha = 0.95$ (Arnett et al., 2011; Burton et al., 2019; Schulz-Zhecheva et al., 2017).

Strengths and Difficulties Questionnaire (SDQ)

The SDQ hyperactivity-inattention subscale (SDQ/HI) was used in this study to compare the SWAN psychometric properties (convergent and concurrent, diagnostic validity) and clinical utility with findings for a widely applied clinical parent-report subscale. The SDQ/HI is a subscale of the SDQ a brief emotional and behavioral screening rating scale. The SDQ involves 25 items, with each set of five items representing each subscale: (1) emotional symptoms, (2) conduct problems, (3) hyperactivity-inattention, (4) peer problems, and (5) prosocial behaviors.

A total difficulty (TD) score is generated from the sum of the first four subscales of emotion, conduct, hyperactivity-inattention, and peer problems (20 items); the items for prosocial behaviors are not included in the total difficulties score. The parent is asked to rate the child's behavior over the past 6 months. Each of the 25 items has the following rating options: not true (0), somewhat true (1), or certainly true (2). Evidence suggests that the SDQ demonstrates good test-retest reliability ($r=.49-.77$), and internal consistency ($\alpha =.58-.85$) (Goodman, 1997; Lakes et al., 2011; Stone, Otten, Engels, Vermulst, & Janssens, 2010). Although the SDQ/HI does not list all the DSM-5 ADHD symptoms, it offers a rating on both of the symptom domains of ADHD: inattention and hyperactivity-impulsivity. Three items of the SDQ/HI are stated as difficulties, including “constantly fidgeting or squirming,” “easily distracted, concentration wanders,” and “restless, overactive, cannot stay still for long”. The two items listed as strengths are “good attention span, sees work through to the end” and “thinks things out before acting”(Goodman, 1997).

For scoring, as suggested by the SDQ guidelines, the two items stated as strengths were reverse scored (2 = not true, 1 = somewhat true, and 0 = certainly true). Previous evidence suggested that the SDQ/HI scale and the TD score performed similarly in predicting ADHD (Algorta, Dodd, Stringaris, & Youngstrom, 2016). A study investigating the SDQ/HI subscale's screening ability in clinical samples of children and adolescents found that a cutoff of 8/10 on the subscale showed the highest diagnostic accuracy for detecting ADHD (Carballo, Rodríguez-Blanco, García-Nieto, & Baca-García, 2014).

The Impairment Rating Scale (IRS)

The Impairment Rating Scale IRS is a measure of ADHD-related functional impairment. The IRS was utilized in this study to examine the concurrent validity of the SWAN compared to

the SDQ/HI. The IRS (Fabiano et al., 2006) is a brief rating scale that assesses a child's degree of impairment across major life domains that are recognized as impaired in children with ADHD. The IRS consists of seven items, and each item represents a specific impairment domain: (a) relationship with peers, (b) relationship with siblings, (c) relationship with parents, (d) academic progress, (e) self-esteem, and (f) influence on family functioning, (g) overall severity of child's problem and child need for treatment. Parents are asked to assess the severity of their child's problem in each domain and the need for treatment or special services. It is a visual-analogue scale; for scoring, a rater is asked to place an X on a line that indicates the severity of a child's problem on an impairment continuum, with anchors ranging from 0 (no problem) to 6 (extreme problem). The line is divided into seven equally spaced segments, and the segment where the X was placed constitutes a score between 0 and 6. A score of 3 or greater indicates significant impairment. In addition, for each domain, there is a section for the rater to describe his or her view of the child's problem in a narrative description.

In the current study, data were collected as a 7-point Likert scale based on parent reports for each item. Shifting the scoring of the IRS from the visual analogue to the Likert scale has been used in previous studies (Fabiano et al., 2006) for ease of data collection and scoring in large sample size studies. The IRS has been validated for use with children and adolescents (Fabiano et al., 2006). Evidence suggests that it has strong test-retest reliability (Fabiano et al., 2006).

Data Analysis

SPSS software version 28.0 was used for the following analyses: preliminary descriptive statistics, *t*, Chi-Square, and Fisher's exact tests when appropriate to compare demographic information (race, ethnicity, sex, and age) between the ADHD and non-ADHD groups, and

bivariate correlation analyses to examine the convergent and concurrent validity among the current study rating scales. Receiver operating characteristic (ROC) curve analyses were used to determine the SWAN and SDQ/HI cutoff scores and compare their performance in identifying cases with ADHD from other disorders in this sample. ROC curve analysis is useful to determine the extent to which a continuous measure can identify clinical ADHD that is defined dichotomously as present or absent using an established reference standard. Generally, the ROC analysis helps identify cutoffs corresponding to balanced levels of scale sensitivity and specificity (Youngstrom, 2014). Graphically, the ROC curve illustrates a plot of true positive rate (sensitivity) along the Y-axis against the false positive rate (1-specificity) along the X-axis for all possible cutoff points of a rating scale score and then combines them into a single value called area under the curve (AUC; Youngstrom, 2014).

The AUC value represents the chance that a randomly selected child with ADHD would score higher on the ADHD rating scale (e.g., SWAN) than a randomly selected child without ADHD. The diagonal line across the ROC curve graph represents a useless ADHD rating scale with an AUC value equal to a null hypothesis ($AUC = 0.5$). If the AUC value for a rating scale is equal to 0.5, then it is not better than chance or a coin flip in detecting cases with ADHD. Conversely, an ADHD rating scale with an AUC of 1.0 means balanced levels of sensitivity and specificity (100% sensitivity and 100% specificity) and, therefore, will better classify cases. Nonetheless, such a rating scale is not realistic. Most likely, a high AUC suggests a study design flaw because a slight failure to separate cases from non-cases is often expected in a real clinical context (Youngstrom, 2014). To evaluate the AUC, Swets's (1988) guideline suggests the following AUC values: 0.60–0.70 = poor; 0.71–0.79 = fair; 0.80–0.89 = good; and 0.90–1.00 = excellent. Also, an additional guideline by Rice and Harris (2005) that previous studies

frequently used as a benchmark to evaluate AUC values based on Cohen's d interprets AUC values as follows: < 0.55 = poor and clinically not useful; $.6-.7$ = sufficient; $.7-.8$ = good; $.8-.9$ = very good; and $> .9$ = excellent (Rice & Harris, 2005). In this study, all ROC curve analyses were based on nonparametric methods and conducted using MedCalc Statistical Software (version 20; MedCalc Software, 2019). The DeLong et al. (1988) method for comparing paired ROC curves was used to compare the AUCs of the SWAN and SDQ. Using the Rice & Harris (2005) guidelines, an AUC value of at least 0.7 was expected to reflect an acceptable ADHD rating scale in this study.

Analysis for Aim 1

Aim 1 was to estimate the SWAN's convergent validity with the SDQ/HI by examining the SWAN total and subscale scores' correlations with the SDQ/HI. To accomplish this, Pearson r correlations between the SWAN subscales, the total score, and the SDQ/HI were estimated. Correlation coefficients between 0.1 and 0.3 were considered low; those between 0.31 and 0.5 were considered moderate; and those over 0.5 were considered high.

Analysis for Aim 2

Aim 2 was to examine the concurrent validity of the SWAN compared to the SDQ/HI, as measured by correlations with scores from the IRS, a measure of ADHD-related functional impairment. To accomplish this, the Pearson r correlations between each of the SWAN subscales, the total score, and the SDQ/HI were estimated with IRS for each item and total score. Correlation coefficients between 0.1 and 0.3 were considered low; those between 0.31 and 0.5 were considered moderate; and those over 0.5 were considered high.

Analysis for Aim 3

Aim 3 was to examine the diagnostic validity of the SWAN by exploring (a) the discriminant abilities of the SWAN total and subscales scores to differentiate between children with and without an ADHD diagnosis and (b) the diagnostic efficiency of the SWAN total and subscale scores to determine the best cutoff scores differentiating children with and without ADHD diagnosis. To accomplish the first part of this aim, ROC curves were plotted for each of the SWAN subscales. The AUC for each rating scale was calculated using clinical ADHD as the reference standard. The critical value for significance for the AUC was $p = .05$. To accomplish the second part of this aim, sensitivity and specificity were calculated. Optimal cutoff scores for the SWAN subscales for differentiating between children with and without ADHD were derived using Youden's index, which is an overall measure of test performance. It was used to identify the optimal cutoff score (best diagnostic cutoff) for the SWAN total and subscales for differentiating between children with and without ADHD.

Analysis for Aim 4

Aim 4 was to compare the discriminant abilities of the SWAN total and subscales with the SDQ/HI scores to differentiate between children with and without an ADHD diagnosis. First, the ROC curve for SDQ/HI subscale was plotted. AUC was calculated to evaluate the SDQ performance in identifying cases of ADHD using clinical ADHD as the reference standard. The critical value for significance for the AUC was $p = .05$. Next, Youden's index was calculated to identify the optimal cutoff score (best diagnostic cutoff) of the SDQ/HI subscale for differentiating between children with and without ADHD. The p values and AUCs of both SWAN scales (total and subscale) and the SDQ/HI were compared to determine which rating scale performed better in differentiating children with ADHD from other children without

ADHD in this sample. The comparison between the p values and AUCs were based on a nonparametric approach that compared the areas of two or more correlated ROC curves (DeLong et al., 1988).

Sample Size

For estimating the sample size for the current study, it was assumed based on results from previous studies that the correlations between the study's ADHD rating scales would be significant and moderate and that the correlation coefficient between the scales would be equal to or greater than 0.5. Thus, it was assumed that at least a sample of 112 participants would be required for this analysis. Regarding AUCs, a priori sample size calculation indicated a requirement of at least 129 subjects, with 43 positive ADHD cases and 86 non-ADHD cases, based on 80% power and applying a significance level of $p < 0.05$. In this study, the sample size for these analyses was above these estimated required samples.

Chapter 4: Results

This chapter presents the results of each aim of the current study. The first aim was to estimate the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) scale's convergent validity with the SDQ hyperactivity-inattention subscale (SDQ/HI) by examining the SWAN total and subscale scores correlations with the SDQ/HI. It was hypothesized that the correlations between the SWAN total and subscales scores and SDQ/HI subscales scores would be both significant and moderate, as evidenced by correlation coefficient equal to or greater than 0.5. The second aim of the study was to compare the concurrent validity of the SWAN with that of the SDQ/HI, as measured by correlations with scores of a measure of ADHD-related functional impairment, the Impairment Rating Scale (IRS). For this aim, it was predicted that the correlations between the SWAN total and subscale scores and the overall IRS score would be both significant and moderate, as evidenced by correlation coefficient equal to or greater than 0.5. The third aim was to examine the diagnostic validity of the SWAN by exploring the following: (a) the discriminant abilities of the SWAN total and subscales scores to differentiate between children with and without an ADHD diagnosis, and (b) the diagnostic efficiency of the SWAN total and subscales scores to determine the best cutoff scores that differentiate between children with and without ADHD diagnosis. The hypothesis for aim three was that the SWAN would discriminate between children with and without ADHD, as evidenced by significant AUC values equal to or greater than 0.7. The final aim of this study was to compare the discriminant abilities of the SWAN total and subscales and SDQ/HI scores to differentiate between children with and without an ADHD diagnosis. It was hypothesized that the SWAN would have greater sensitivity and specificity to differentiate ADHD than the SDQ/HI.

Sample

This study utilized data of 357 children aged 6–11 years who visited a mental health outpatient clinic specializing in ADHD and related disorders, at a children's hospital.

All study participants were referred from community pediatricians, schools, and mental health professionals between January 1, 2015, and August 31, 2019. A total of 257 children were ADHD cases, and 100 were non-ADHD (children with other psychiatric or medical diagnoses). Most of the sample were male subjects ($n = 262$), and slightly over half (53%) of the male subjects were ADHD cases. Female subjects represented somewhat more than a quarter of this sample ($n = 95$), and only 19% of them were ADHD cases. Regarding the race of the subjects, the majority were White, or Caucasian children (65%); Asian children represented 11%; and Black, or African American, children represented 5% of the sample. No differences between the ADHD group and the non-ADHD group for age, sex, race, and ethnicity were found as indicated by t , Chi-Square, and Fisher's exact tests. For age, Levene's test for equality of variances was conducted and was not significant ($p > .05$), suggesting that that the two groups were comparable.

Descriptive Statistics

Values for mean, standard deviation, and range of scores for the SWAN, SDQ/HI and each of the IRS items are presented in Table 4.1. The Kolmogorov–Smirnov test, and the visual inspection of histograms revealed that the SWAN total scores were normally distributed in the current sample. However, the SWAN subscales as well as the SDQ/HI were not normally distributed.

Study Results

Findings are listed by aim and hypothesis as follows:

Convergent Validity Findings

Aim 1: The first aim of this study was to estimate the SWAN's convergent validity with the SDQ/HI by examining the SWAN total and subscale scores' correlations with the SDQ/HI. The convergent validity of the SWAN total and the subscales scores was examined by calculating the bivariate Pearson r correlation coefficients among the raw scores of the SWAN total, the SWAN Hyperactivity/Impulsivity subscale (SWAN HI) subscale, and the SWAN Attention Deficit subscale (SWAN AD) subscale with the SDQ /HI subscale. Table 4.2 illustrates the correlation matrix of all scales.

Results for Hypothesis 1. It was predicted that the correlations among the SWAN total, the SWAN HI, the SWAN AD scores, and the SDQ/HI subscales scores would be significant and moderate, as evidenced by the correlation coefficient equal to or greater than .5. As illustrated by Table 4.2, the SWAN total and SDQ/HI correlation were both significant and moderate ($r = .69$). Likewise, the correlation between the SDQ/HI and the SWAN HI was significant and moderate ($r = .67$). However, the correlation between the SDQ/HI and the SWAN AD subscale was significant ($r = .48$) but smaller than the correlation magnitudes between SDQ/HI with SWAN HI and SDQ/ HI with SWAN total. Scatter plots were used to visually illustrate the SWAN total and SDQ/HI relationship (see Figure 4.1).

Concurrent Validity

Aim 2: The second aim of the study was to examine the concurrent validity of the SWAN compared to the SDQ/HI as measured by correlations with scores of a measure of ADHD-related functional impairment, the Impairment Rating Scale (IRS). The concurrent validity was examined by calculating the bivariate Pearson's r correlation coefficients between each item of

the IRS scale and the ADHD rating scales: the SWAN total and the SWAN HI subscale, the SWAN AD subscale, and the SDQ/I subscale (see Table 4.3).

Results for Hypothesis 2. It was hypothesized that the SWAN total and the SWAN subscales would have larger correlations with the IRS items than the SDQ/HI, as evidenced by the significant and moderate correlation coefficient equal to or greater than .5. As indicated in Table 4.3, for concurrent validity, the correlation coefficient of the SWAN total and the SWAN subscales and the SDQ/HI with the IRS items were all small. The SWAN total compared to the SWAN AD, SWAN HI and SDQ/HI subscales in this study had the largest correlation with the IRS overall impairment item score; however, it was relatively small ($r = .31$). Similarly, the correlation coefficients of the IRS overall impairment score with SWAN HI, SWAN AD, and SDQ/HI were significant and low (less than $r=.3$). The SWAN HI subscale had the second highest correlation value with the IRS overall item ($r =.28$), followed by the SWAN AD subscale ($r =.24$); the SDQ/HI had the lowest correlation value with the IRS overall impairment ($r =.23$).

The SWAN total score correlations with other IRS items were low. The highest correlation with the other IRS items was in the academic progress impairment ($r =.36$), followed by the impairments in relationships with playmates and general impairment in the family; both correlations were small and equal to one another ($r =.23$); finally, correlation with impairment in the relationship with parents was small as well ($r =.19$). Likewise, the SWAN HI subscale correlations with the other IRS items were small and suggested that children demonstrated greater impairments in the relationship with playmates ($r =.3$), followed by general impairment in the family, relationship with parents, and academic progress (r values ranged from .19 to .23). The SWAN AD subscale correlations with the IRS items were significant with only two items: greater impairment in academic progress ($r =.41$) and lower-level general impairment in the

family ($r = .14$). Overall, the SDQ/HI had lower correlation coefficient values with the IRS items than the SWAN total and its subscales. All significant correlations of the SDQ/HI with the IRS items were consistently low (less than $r = .3$). Notably, none of the SWAN total, SWAN subscales or the SDQ/HI showed significant associations with parent-reported impairment in child self-esteem and in relationships with siblings. However, it should be noted that impairment with sibling response was only required if applicable (i.e., the child had siblings); thus, there were missing values of ($n = 98$), and pairwise deletion was used to maximize the available sample size.

Diagnostic Validity

Aim 3: The third aim of the study was to examine the diagnostic validity of the SWAN. Receiver operating characteristic (ROC) analyses were used to calculate the following: (a) the Area Under Curve (AUC), and (b) the optimum cutoff scores of the SWAN total and the subscale to determine differentiating children with and without ADHD diagnosis.

Results for Hypothesis 3. It was predicted that the SWAN would discriminate between children with and without ADHD, as evidenced by significant AUC values equal to or greater than 0.7. Table 4.4 and Figure 4.2 present the ROC curves and AUC for each of the SWAN subscales. Based on the Rice and Harris guidelines (2005), the SWAN's total ability to differentiate between children with and without an ADHD diagnosis was efficient (AUC = .71). However, the discriminant abilities of the SWAN AD and the SWAN HI were slightly lower than the SWAN total. The SWAN AD and SWAN HI both demonstrated adequate ability (AUC = 0.69 and 0.66, respectively) in differentiating between children with and without an ADHD diagnosis. The results from the pairwise comparison of the ROC curves with the critical value for significance for the AUC was set at ($p = .05$) and showed no significant differences between the AUCs of the

SWAN total and the SWAN AD. However, the difference between the SWAN total and the SWAN HI AUC was significant ($p = 0.023$).

Table 4.4. shows the optimal cutoff scores, sensitivity, and specificity based on Youden's index for each of the SWAN totals and subscales. For the SWAN total, a score of > 0.72 (sensitivity = 83%, specificity = 47%) was the optimum score for identifying ADHD in this study sample. The optimum cutoff score for the SWAN AD was > 0.66 (sensitivity = 82.7%, specificity = 48.6%) and for the SWAN HI was a score of > 0.77 (sensitivity = 82.7%, specificity = 49.5%).

Comparing Diagnostic Validity of SWAN with SDQ/HI

Aim 4: The fourth aim of this study was to compare the discriminant abilities of the SWAN total and the subscales with the SDQ/HI scores to differentiate between children with and without an ADHD diagnosis. An ROC analysis was used to calculate the AUC and the optimum cutoff score of the SDQ/HI that differentiated children with and without ADHD diagnoses. The ROC curve analysis for the SDQ /HI, as shown in Table 4.4., indicated that the SDQ/HI had adequate discrimination (AUC = 0.69) when differentiating between children with and without an ADHD diagnosis. The SDQ/HI optimum cutoff based on Youden's index was > 6 (sensitivity = 77.9 %, specificity = 49.6%). Figure 4.3 presents the ROC curves for the SDQ/HI when differentiating between the ADHD and non-ADHD groups.

Results for Hypothesis 4. It was hypothesized that the SWAN would have greater sensitivity and specificity in identifying children with ADHD than the SDQ/HI. Figure 4.4 presents the ROC curves for the SWAN total and the SWAN HI subscale as well as the SWAN AD subscale with the SDQ/HI subscale when differentiating between the ADHD and non-ADHD groups. As shown in Figure 4.4, the SWAN total had the largest AUC among all scales. However, the results

from the pairwise comparison of the ROC curves with the critical value for significance for the AUC was set at $p = .05$ and showed no significant differences between the AUCs of the SWAN total, the SWAN HI subscale, and the SWAN AD subscale with the SDQ/HI. These results suggest no significant difference between the abilities of the SWAN and the SDQ/HI to differentiate between the ADHD and non-ADHD groups.

Chapter 5: Discussion

The SWAN is a novel ADHD behavioral rating scale that identifies children's strengths and weaknesses along the full ADHD continuum, not just a weakness or problem. Although several studies used the SWAN in community settings, little is known about its utility in a clinic population as a diagnostic measure for identifying cases of ADHD (Brites et al., 2015; Collett et al., 2003). The few studies that examined the SWAN's ability to identify cases with ADHD in clinically referred children had the limitation of using a criterion reference based solely on a structured parent interview (Chan et al., 2014; Robaey et al., 2007). To our knowledge, no published study has evaluated the SWAN ability to classify ADHD and differentiate it from other (non-ADHD) diagnoses using the best practices of ADHD assessment as a standard reference. This study sought to investigate the SWAN's convergent validity by examining its association with a common ADHD rating subscale; Hyperactivity–Inattention (HI) of the Strengths and Difficulties Questionnaire (SDQ/HI). The second aim of the study was to compare the concurrent validity of the SWAN with that of the SDQ/HI, as measured by correlations with scores on a measure of ADHD-related functional impairment, the Impairment Rating Scale (IRS). Focusing on the usefulness of the SWAN for diagnostic evaluation in a clinical context, this study evaluated the SWAN's total and subscale's ability to classify ADHD and differentiate it from other (non-ADHD) diagnoses compared to SDQ/HI. Parents of children referred for ADHD evaluation in an outpatient clinic specializing in ADHD and related disorders at a children's hospital completed the SWAN as part of a pre-visit online routine. All participants received a comprehensive clinical evaluation to determine their diagnoses. Overall, the findings of this study provided mixed evidence to support the usefulness of the SWAN for a diagnostic evaluation in clinical settings.

Convergent Validity

The first aim of this study was to examine how the SWAN total and subscales correlated with SDQ/HI. It was hypothesized that the correlations between the SWAN total and subscales scores and SDQ/HI subscales scores would be both significant and moderate, as evidenced by correlation coefficients equal to or greater than 0.5. Consistent with what was hypothesized, the results of this study showed the SWAN total, and the SWAN Hyperactivity/Impulsivity subscale (SWAN HI) had a high and significant correlations with SDQ/HI ($r = .68$, $r = .65$) respectively. Similarly, the correlation between SWAN Attention Deficit subscale (SWAN AD) and the SDQ/HI was significant and moderate ($r = .48$). These results support the idea that both rating scales are measuring a similar underlying construct.

These results are consistent with findings from a study by Lakes et al. (2011) that examined the convergent validity of the SWAN total and subscales with SDQ based on data of preschool children aged 3–5 years whose parents participated in a community parent education program. They found evidence to support convergent validity between the SWAN total and SDQ/HI based on a significant correlation ($r = .54$). Further, they found the correlations between the SWAN AD, SWAN HI with SDQ/HI were ($r = .52$, $r = .49$) respectively

Concurrent Validity

The second aim of this study was to compare the concurrent validity of the SWAN with that of SDQ/HI, as measured by correlations with scores from the Impairment Rating Scale (IRS), a measure of ADHD-related functional impairment. It was hypothesized that the SWAN total and the SWAN subscales would have larger correlations with the IRS items than with the SDQ/HI, as evidenced by the significant and moderate correlation coefficient equal to or greater than 0.5. Consistent with our hypotheses, the results revealed higher correlations between the

IRS items with the SWAN total than with SDQ/HI. These findings were expected given that the SWAN has a broader content—18 items aligned with DSM 5 ADHD criteria—whereas SDQ/HI is part of a broadband scale with only five items that tap into ADHD. Among the ADHD rating scales in this study, the SWAN total showed the highest correlation with the IRS item “overall impairment.” The SWAN HI demonstrated the highest agreement with other IRS items, including impairment in playing with other children and impairment in relationships with parents and family in general. The magnitude of correlations between the IRS items and the SWAN AD was smaller than between the IRS items and the SWAN total and SWAN HI. However, the SWAN AD had the strongest association with the IRS academic impairment item ($r = .41$). These findings align with previous studies wherein inattentive symptoms strongly correlated with poor school performance (Abrines et al., 2012; Garner et al., 2013b; Rigoni, Blevins, Rettew, & Kasehagen, 2020 & Kasehagen, 2020). In contrast, hyperactivity and impulsivity symptoms are largely associated with maladjustments in many social relationships such as parent and peer relationships (Graziano, McNamara, Geffken, & Reid, 2011 & Reid, 2011; Zoromski, Epstein, & Ciesielski, 2021 2021).

Nevertheless, the magnitude of correlations between the IRS items and the SWAN total and subscales was smaller than expected. In general, it is well documented in the literature that ADHD symptoms and functional impairment ratings have modest correlations (Gathje, Lewandowski, & Gordon, 2008 2008; Sasser et al., 2017 2017). Gordon et al. (2006) found modest correlations between rating scales of ADHD symptoms and ratings of functional impairments that never exceeded a sample correlation coefficient of .43. They found this result was consistent among four large-scale ADHD research samples of children ages 3–18 years. Thus, the low to moderate correlations between the SWAN and the IRS in the current study are

consistent with research showing ADHD symptoms alone do not necessarily indicate ADHD diagnosis. A child may have many ADHD symptoms without showing any maladjustments, whereas another child with fewer ADHD symptoms (i.e., subthreshold ADHD) may experience significant impairments in different life realms. That is to say, ADHD diagnosis requires a thorough assessment of functional impairment - more information than what is obtained by a symptoms rating scale. Evaluation of functional impairment is crucial for not just accurate ADHD diagnosis but also guiding and monitoring the treatment outcomes.

The impact of the restricted range in this sample is an additional potential justification for the low to moderate correlations between the SWAN and the IRS items. This sample presented with a high rate of ADHD prevalence, over 70% were ADHD cases. The correlations between the ADHD symptoms measured by the SWAN and the IRS scores would be much higher if non-ADHD individuals were included.

An additional alternative explanation for these low to moderate correlations could be the features of the IRS scale. The IRS is a single item per domain measure where each item allows parents to report impairments in one life domain. The brevity of the IRS might be insufficient to adequately capture nuances of a child's impairments (Sibley et al., 2017). An added limitation of the IRS is it asks parents to rate a child's impairment levels and the child's need for service with single questions. Therefore, parents must answer compound questions that touch separate issues with a single response. Although the scale's author argued this combination was to avoid unnecessary questions (Fabiano et al., 2006), simultaneously rating a child's impairment level and the child's need for services might confuse parents. An additional issue with IRS that might lead to respondent confusion is that the IRS instructions provide no requirements that the impairment is primarily related to ADHD (Sibley et al., 2012). Thus, future studies examining

SWAN score correlations with a more comprehensive ADHD specific impairment scale are warranted to further investigate the SWAN's concurrent validity.

Diagnostic Validity

It was predicted that the SWAN would discriminate between children with and without ADHD, as evidenced by significant area under the curve (AUC) value equal to or greater than (AUC=.7). In this study we used the Rice and Harris (2005) guidelines that previous studies frequently used as a benchmark to evaluate AUC values based on Cohen's *d*; interpretation of AUC values was as follows: <.55 = poor and clinically not useful; .6-.7 = sufficient; .7-.8 = good; .8-.9 = very good; and > .9 = excellent (Rice & Harris, 2005). In this study, the SWAN total and subscales showed good to sufficient abilities to differentiate between children with ADHD and those with a wide variety of other possible clinical disorders as evidenced by AUCs for SWAN total, SWAN AD, and SWAN HI (AUCs = .71, =.69 and .66) respectively. Based on the Rice and Harris guidelines (2005), these findings provide mixed evidence to support our hypothesis. The SWAN total showed good ability (AUC = .71) to discriminate between children with and without ADHD. The SWAN AD and SWAN HI were slightly less accurate than the SWAN total, demonstrating sufficient ability to distinguish between children with and without ADHD. The results from the pairwise comparison of the ROC curves with the critical value for significance for the AUC was set at ($p = .05$) showed no significant differences between the AUCs of the SWAN total and the SWAN AD. However, the difference between the SWAN total and the SWAN HI AUCs was significant ($p = .023$), suggesting that the SWAN total and SWAN AD were more accurate than the SWAN HI in discriminating ADHD from non-ADHD children.

The optimum cutoff scores were derived empirically in this sample by applying Youden's index, which represented the best balance score that maximizes both sensitivity and specificity,

demonstrated by the SWAN differentiating between children with and without ADHD with high sensitivity and low specificity. These findings suggest the SWAN is clinically useful for screening. The SWAN is unlikely to miss children with possible ADHD who need further evaluation. However, when the goal of the assessment is a diagnostic evaluation to rule out other disorders, the SWAN total and subscales had limited capacity to do so in this sample. These results are inconsistent with studies that established SWAN cutoff scores with a balanced trade-off between sensitivity and specificity to differentiate between children with ADHD and children with other psychiatric disorders. Findings from a study by Robaey et al. (2007), in a sample of children aged 6–9 years who were referred to a clinic for ADHD evaluation showed the French-SWAN total score had high sensitivity (86%) and specificity (88%). However, the results of their study were limited by a small sample size of 74 children with ADHD and 14 without an ADHD diagnosis.

Similarly, a study by Chan et al. (2014) found the Chinese-SWAN subscales (SWAN AD and SWAN HI) effectively differentiated ADHD from non-ADHD psychiatric controls among 290 children ages 6–12 years in a psychiatric clinic setting. Both SWAN AD and SWAN HI showed reasonable to excellent sensitivity and specificity, ranging from 91% to 96%, and 74% to 95%, respectively (Chan et al., 2014). Overall, these two studies used structured parent interviews (DISC-IV) as references for ADHD diagnosis, which might increase the risk of shared method variance between the SWAN scores and clinical diagnosis because both scales were obtained from parents. Conversely, the current study defined the criterion diagnosis for ADHD based on a comprehensive clinical evaluation of ADHD from multiple informants. Thus, results from these two studies might overestimate the SWAN's diagnostic validity, unlike the current study.

The referral pattern in this sample could further clarify the lower discriminant abilities of the SWAN in this study compared to previous studies. Most participants in the current study were referred by community pediatricians, schools, and mental health professionals to a clinic that specialized in ADHD and related disorders for further evaluation when ADHD diagnosis was not clear. This study's sample had many complex cases with conditions that could mimic ADHD symptoms and comorbidities (e.g., speech and language developmental disorders, anxiety). Consequently, it is not surprising that the SWAN scores in this sample were only sufficient (AUC=.71) but not near excellent or very good (AUC greater than .71) in discriminating children with ADHD, where more information beyond that provided by a rating scale was needed to arrive at a diagnosis.

Comparing Diagnostic Validity of SWAN with SDQ/HI

The final aim of this study was to compare the SWAN's discriminant ability to that of the SDQ/HI. SDQ/HI showed a sufficient ability to classify ADHD and differentiate it from other psychiatric, non-ADHD diagnoses as evidenced by AUC of .69. The SDQ/HI optimum score (> 6) yielded a sensitivity of 77.9% and a specificity of 49.5%. This finding is contrary to findings in a study by Carballo, Rodríguez-Blanco, García-Nieto, and Baca-García (2014) that investigated the screening ability of SDQ/HI for ADHD in a clinical sample of children aged 3–17 years. Carballo et al. (2014) found the SDQ/HI score of 6 yielded a sensitivity of 70% and specificity of 84%.

The discrepancy between this study's findings and those in Carballo et al. (2014) could be owing to their sample having a wider age range. Further, unlike this study, which used a comprehensive clinical assessment, their gold standard was based on a parent report rating scale,

the ADHD Rating Scale-IV (DuPaul et al., 1998). Thus, their results might overestimate SDQ/HI owing to shared method variance because the same parents completed both scales.

It was hypothesized that the SWAN would have greater ability in identifying children with ADHD than the SDQ/HI. In agreement with this hypothesis, the findings from the current study suggest that the SWAN total score ($AUC = .71$) was better able than SDQ/HI ($AUC = .69$) to discriminate ADHD cases in the present sample. However, the results from the pairwise comparison of the ROC curves revealed that there was not a significant difference between the AUCs of the SWAN and SDQ/HI. In other words, there was no significant difference between the two scales in identifying ADHD cases in this sample. This finding is somewhat surprising because SDQ/HI is a short subscale with only five items, whereas the SWAN is an ADHD-specific scale with items based on the DSM's ADHD symptoms.

In light of these findings, we can conclude that each of the SWAN and the SDQ/HI can be useful for different clinical purposes. The SDQ/HI might be a better screening tool for several reasons. First, in a clinical context, the brevity of SDQ/HI is an advantage. SDQ/HI would be easier to be completed by parents and scored by clinicians. Second, similar to the SWAN, SDQ/HI is a strength-based rating scale highlighting a child's strengths and problems. Third, SDQ/HI is part of the SDQ rating scale, a broad and multidimensional rating scale that allows screening for comorbid symptoms and ADHD symptoms. In addition to SDQ/HI, the SDQ contains the following four subscales: emotional symptoms, conduct behaviors, peer problems, and prosocial behavior. Finally, U.S. normative data for the SDQ are available based on age and sex (Bourdon, Goodman, Rae, Simpson, & Koretz, 2005; Simpson, & Koretz, 2005), but such data are still lacking for the SWAN (Arnett et al., 2011). Thus, the SDQ scale appears to be preferable over the SWAN for screening. Overall, it should be noted both SDQ/HI and the

SWAN had low specificity (49.5%, 46.5, respectively) in the current sample. Sensitivity and specificity are equally crucial for an ADHD rating scale and valuing one over the other is not recommended (Lai et al., 2011; Parker & Corkum, 2016).

Although the SWAN had relatively comparable abilities with SDQ/HI in identifying cases with ADHD in this ADHD clinical sample, variations about strengths in attention and activity abilities provide additional information for clinical and research purposes.

Limitations and Strengths

When interpreting the findings of the current study it is important to consider the following limitations. The current study had a high risk of diagnostic review bias because clinicians at the study site were not blinded to the SWAN and SDQ/HI results when evaluating the children. This bias occurs when there are criterion contaminations that make the outcome (the clinical ADHD diagnosis) not independent of the predictor (SWAN and SDQ/HI scores). Such bias is an inevitable risk associated with retrospective study designs that use existing archived real-world clinical data. Nonetheless, the effects of this bias on this study's findings are likely to be minimal, as in most cases, the SWAN was not scored before or during the clinic visit and was not in the patients' medical charts. The current study results revealed the SWAN had high sensitivity and poor specificity (high false positives). This result suggests imperfect agreement between the SWAN elevated scores and clinical diagnoses. The risk of an inflated relationship between SWAN scores and ADHD clinical diagnoses would have been higher in the current study if clinicians relied on the SWAN scores. However, it seemed clinicians in the present study were not relying heavily on the SWAN to inform their diagnosis.

An additional limitation is the shared method variance. The study rating scales, and semi-structured interview sections of children's evaluations were both based on parent reports.

However, clinical diagnosis in the current study was informed by multiple measures from multiple informants, including both teachers and parent ratings and clinical observations.

Although the use of a sample of clinically referred children is a strength of this study, this could be considered a limitation as well. The current study's findings are specific to a sample from an ADHD tertiary clinic, wherein the majority of referred cases were complex with significant psychiatric comorbidity and diverse medical and treatment history. Hence, the present study's findings may not extend to other settings such as primary health care, pediatric clinics, and school settings. Additionally, this sample was predominantly White, non-Hispanic, with limited ethnic diversity. As such, the results may not be generalizable to other ethnic groups.

Despite these limitations, this study was strengthened by using a large sample of children referred to a clinic specializing in ADHD who received a comprehensive multidisciplinary evaluation. An additional strength of the study was using a diagnostic standard for ADHD clinical diagnosis based on the best practices of ADHD assessment. The quality of a diagnostic validity study is linked to the quality of the diagnostic standard. Overall, from a theoretical perspective, this inquiry provided an understanding of the extent to which ADHD symptoms—conceptualized dimensionally in the SWAN as the end of behavioral dimensions normally distributed in the population—converge with clinical ADHD diagnosis conceptualized as a categorical disorder.

Future Research Direction

Many questions remain unanswered about the SWAN's clinical utility. The results of this study addressed only some of the gaps in the current literature about the SWAN's validity in a sample of children with ADHD comorbidities and other psychiatric disorders. Future research should seek to replicate the findings of this study in non-clinical settings such as primary health

care and school settings which would likely improve the screening capacity of the SWAN if there were more typically developing children in such samples. Further, other SWAN psychometric properties have not yet been investigated in clinical samples. This includes test-retest reliability and comparing the diagnostic validity of the SWAN's parent format and teacher format, which has been investigated using only the Chinese version (Chan et al., 2014). An additional promising line of research for longitudinal studies is to examine the stability of the SWAN scores as children mature into adolescents.

Clinical Implications

Results from this study have some important clinical implications. The findings of this study suggest the SWAN scale exhibited good ability in discriminating children with ADHD children from children without ADHD in clinic-referred children. While the SWAN demonstrated low specificity in the current study, that was not surprising for a sample of children clinically referred for attention problems. The SWAN would have presented better discriminating abilities in the general population or samples with fewer non-ADHD cases.

Although the SWAN had relatively equivalent abilities with SDQ/HI in identifying cases with ADHD in this ADHD clinical sample, variations (strengths) in attention and activity abilities provide additional information for clinical and research purposes. The SWAN scores generated a near-normal distribution in the current clinical referred sample where at least each participant had one psychiatric disorder. Thus, our findings suggest advantages of the SWAN over SDQ/HI for describing unique aspects of clinic-referred children, that may be relevant for predicting outcome and treatment planning.

The value of the added information of the SWAN to the clinical practice is still not well studied. Most of studies conducted in community and population samples suggested that a child's

SWAN scores above the average (e.g., outstanding attention skills) may indicate adaptive outcomes. For instance, Kiani, Hadianfard, and Mitchell (2019), in a community sample of 280 Iranian children aged 6–12 years, conducted group analysis to compare functional impairment outcomes, quality of life, and emotional and behavioral problems among children with strong, moderate, and weak attention and behavioral control based on the SWAN scores. The study revealed that the children in the group with strong attention and control skills (i.e., strengths) had a better quality of life, physical, emotional functioning, school, and social functioning and showed fewer conduct problems (Kiani et al., 2019)

Nevertheless, it remains unclear whether these strengths convey advantageous or adverse characteristics in children in clinical populations. It has been recommended that future studies could explore outcomes associated with strengths at each SWAN item distribution (Alexander, Salum, Swanson, & Milham, 2020a). A unique feature of the SWAN is the opportunity to explore individual items' strengths despite a child's SWAN total scores. A clinician can still highlight a child's strengths per item if present. Such information can be helpful for a clinician to understand further whether this strength conveys compensation skill, protective characteristics, or even a sign of impairment. For example, a child's total scores might indicate subthreshold ADHD, yet he/she showed some strength on some of the SWAN items (e.g., listening when being spoken to or playing quietly). Results from studies that attempted to understand strategies among girls with ADHD revealed that compliance and socially desirable behavior might be compensatory behaviors that mask disorders in this group. Therefore, this group might be underdiagnosed because parents usually rate them as less impaired. SWAN strengths scores might be used to examine if the child is trying to compensate for ADHD symptoms. An additional example is when a child's SWAN scores indicate sound organizational skills (e.g.,

organizing her/his papers for schoolwork), yet he/she avoids starting on demanding tasks such as homework and reading. Organizational skills training, a common intervention for teens with ADHD, would not be appropriate in this case. At the same time, if a clinician highlights this child's strength (e.g., organization), it might be utilized to explore possible protective factors (e.g., organized home environment).

The SWAN strength-focused approach is likely to be acceptable to caregivers as it reduces stigma. Sharing information about strengths with the child and family might also increase engagement and self-esteem. The SWAN also has potential applications for other disorders where the full range of attention and activity level may have relevance, such as autism spectrum disorder or mood disorders. Indeed, findings from our study suggest that the SWAN offers an assessment tool that captures strengths even in a clinic-referred sample with frequent comorbidity and range of severity.

The development of more representative population norms for the SWAN would be helpful for clinicians to compare a child's scores with children of the same sex and age. Future studies that will examine the outcomes associated with extra information provided by each of the SWAN items distribution could inform clinicians to develop an individualized treatment plan that addresses child strengths and problems that otherwise be overlooked by other rating scales.

Conclusion

In conclusion, with limited data about the SWAN's diagnostic accuracy in clinical samples, the current study is one of the few studies that adds to the literature on the SWAN's clinical utility. The current study's findings suggest that the SWAN showed a good ability for classifying ADHD cases from non-ADHD cases in clinic-referred children with other psychiatric disorders. Results from comparing the diagnostic validity of the SWAN to SDQ/HI yielded no

significant difference in their performance in identifying ADHD cases in the current sample. Nonetheless, the SWAN has advantages over SDQ/HI for describing aspects of variations (strengths) in attention and activity abilities that may be relevant for predicting outcomes and treatment planning. These are new and noteworthy findings, given that no previous studies had compared the clinical utility of the SWAN and SDQ/HI. While the SWAN demonstrated low specificity in the current study, these results might be specific to the current sample and not be generalized to other samples with different characteristics. Finally, it should be considered that the SWAN was developed to screen for ADHD in the general population, wherein ADHD is less frequent, unlike in this study's clinically referred sample, where 73% of the sample met the criteria for ADHD. Further work is certainly required to examine the SWAN's clinical utility in clinical samples with large groups of non-ADHD patients, such as those with autism spectrum disorder or mood disorders.

References

- Abrines, N., Barcons, N., Brun, C., Marre, D., Sartini, C., & Fumadó, V. (2012). Comparing ADHD symptom levels in children adopted from Eastern Europe and from other regions: Discussing possible factors involved. *Children and Youth Services Review, 34*(9), 1903-1908. doi:10.1016/j.chidyouth.2012.05.025
- Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs: General and Applied, 80*(7), 1-37. doi:10.1037/h0093906
- Alexander, L. M., Salum, G. A., Swanson, J. M., & Milham, M. P. (2020a). Measuring strengths and weaknesses in dimensional psychiatry. *Journal of Child Psychology and Psychiatry, 61*(1), 40-50. doi:10.1111/jcpp.13104
- Alexander, L. M., Salum, G. A., Swanson, J. M., & Milham, M. P. (2020b). Measuring strengths and weaknesses in dimensional psychiatry. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 61*(1), 40-50. doi:10.1111/jcpp.13104
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.
- Arnett, A. B., Pennington, B. F., Friend, A., Willcutt, E. G., Byrne, B., Samuelsson, S., et al. (2011). The SWAN captures variance at the negative and positive ends of the ADHD symptom dimension. *Journal of Attention Disorders, 17*(2), 152-162. doi:10.1177/1087054711427399
- Arnold, L. E., Hodgkins, P., Caci, H., Kahle, J., & Young, S. (2015). Effect of treatment modality on long-term outcomes in attention-deficit/hyperactivity disorder: a systematic review. *PloS one, 10*(2), e0116407-e0116407. doi:10.1371/journal.pone.0116407

- Asherson, P., & Trzaskowski, M. (2015). Attention-deficit/hyperactivity disorder is the extreme and impairing tail of a continuum. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(4), 249-250. Retrieved from <https://doi.org/10.1016/j.jaac.2015.01.014>. doi:10.1016/j.jaac.2015.01.014
- Barbarese, W. J. (2020). The SDBP complex ADHD clinical practice guideline: It's about time. *Journal of Developmental & Behavioral Pediatrics, 41*. Retrieved from https://journals.lww.com/jrnldb/Fulltext/2020/03001/The_SDBP_Complex_ADHD_Clinical_Practice_Guideline_.2.aspx.
- Boland, H., DiSalvo, M., Fried, R., Woodworth, K. Y., Wilens, T., Faraone, S. V., et al. (2020). A literature review and meta-analysis on the effects of ADHD medications on functional outcomes. *Journal of Psychiatric Research, 123*, 21-30. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0022395619314116>. doi:<https://doi.org/10.1016/j.jpsychires.2020.01.006>
- Bosch, A., Bierens, M., de Wit, A. G., Ly, V., van der Velde, J., de Boer, H., et al. (2020). A two arm randomized controlled trial comparing the short and long term effects of an elimination diet and a healthy diet in children with ADHD (TRACE study). Rationale, study design and methods. *BMC Psychiatry, 20*(1), 262. doi:10.1186/s12888-020-02576-2
- Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The Strengths and difficulties questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(6), 557-564. Retrieved from <https://doi.org/10.1097/01.chi.0000159157.57075.c8>. doi:10.1097/01.chi.0000159157.57075.c8

- Brites, C., Salgado-Azoni, C. A., Ferreira, T. L., Lima, R. F., & Ciasca, S. M. (2015). Development and applications of the SWAN rating scale for assessment of attention deficit hyperactivity disorder: A literature review. *Brazilian Journal of Medical and Biological Research = Revista Brasileira de Pesquisas Medicas e Biologicas*, 48(11), 965-972. doi:10.1590/1414-431x20154528
- Brown, T. A., & Barlow, D. H. (2005). Dimensional versus categorical classification of mental disorders in the fifth edition of the diagnostic and statistical manual of mental disorders and beyond: Comment on the special section. *Journal of Abnormal Psychology*, 114(4), 551.
- Burton, C. L., Wright, L., Shan, J., Xiao, B., Dupuis, A., Goodale, T., et al. (2019). SWAN scale for ADHD trait-based genetic research: A validity and polygenic risk study. *Journal of Child Psychology and Psychiatry*, 0(0). doi:10.1111/jcpp.13032
- Carballo, J. J., Rodríguez-Blanco, L., García-Nieto, R., & Baca-García, E. (2014). Screening for the ADHD phenotype using the strengths and difficulties questionnaire in a clinical sample of newly referred children and adolescents. *Journal of Attention Disorders*, 22(11), 1032-1039. Retrieved from <https://doi.org/10.1177/1087054714561858>. doi:10.1177/1087054714561858
- Carlson, K. D., & Herdman, A. O. (2010). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, 15(1), 17-32. doi:10.1177/1094428110392383
- Chan, G. F.-C., Lai, K. Y.-C., Luk, E. S.-L., Hung, S.-F., & Leung, P. W.-L. (2014). Clinical utility of the Chinese Strengths and Weaknesses of ADHD-Symptoms and Normal-

- Behaviors questionnaire (SWAN) when compared with DISC-IV. *Neuropsychiatric Disease and Treatment*, *10*, 1533-1542. doi:10.2147/NDT.S65879
- Chang, L.-Y., Wang, M.-Y., & Tsai, P.-S. (2016). Diagnostic accuracy of rating scales for attention-deficit/hyperactivity disorder: A meta-analysis. *Pediatrics*, *137*(3), e20152749. doi:10.1542/peds.2015-2749
- Chang, Z., Ghirardi, L., Quinn, P. D., Asherson, P., D'Onofrio, B. M., & Larsson, H. (2019). Risks and benefits of attention-deficit/hyperactivity disorder medication on behavioral and neuropsychiatric outcomes: A qualitative review of pharmacoepidemiology studies using linked prescription databases. *Biological Psychiatry*, *86*(5), 335-343. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0006322319312740>. doi:<https://doi.org/10.1016/j.biopsych.2019.04.009>
- Chang, Z., Lichtenstein, P., D'Onofrio, B. M., Sjölander, A., & Larsson, H. (2014). Serious transport accidents in adults with attention-deficit/hyperactivity disorder and the effect of medication: A population-based study. *JAMA Psychiatry*, *71*(3), 319-325. Retrieved from <https://doi.org/10.1001/jamapsychiatry.2013.4174>.
- Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest*, *18*(2), 72-145. Retrieved from <https://doi.org/10.1177/1529100617727266>. doi:10.1177/1529100617727266
- Coghill, D., & Sonuga-Barke, E. J. S. (2012). Annual research review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental

- disorders – implications of recent empirical study. *Journal of Child Psychology and Psychiatry*, 53(5), 469-489. doi:10.1111/j.1469-7610.2011.02511.x
- Cohen, J. F., Korevaar, D. A., Altman, D. G., Bruns, D. E., Gatsonis, C. A., Hooft, L., et al. (2016). STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ open*, 6(11), e012799-e012799. doi:10.1136/bmjopen-2016-012799
- Collett, B. R., Ohan, J. L., & Myers, K. M. (2003). Ten-year review of rating scales. v: scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(9), 1015-1037. Retrieved from <https://doi.org/10.1097/01.CHI.0000070245.24125.B6>. doi:10.1097/01.CHI.0000070245.24125.B6
- Danielson, M. L., Bitsko, R. H., Ghandour, R. M., Holbrook, J. R., Kogan, M. D., & Blumberg, S. J. (2018). Prevalence of parent-reported adhd diagnosis and associated treatment among U.S. children and adolescents, 2016. *Journal of Clinical Child & Adolescent Psychology*, 47(2), 199-212. Retrieved from <https://doi.org/10.1080/15374416.2017.1417860>. doi:10.1080/15374416.2017.1417860
- Epstein, J. N., & Loren, R. E. A. (2013). Changes in the definition of ADHD in DSM-5: Subtle but important. *Neuropsychiatry*, 3(5), 455-458. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/24644516> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3955126/>. doi:10.2217/npv.13.59
- Fabiano, G. A., Pelham, J. W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., et al. (2006). A practical measure of impairment: Psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child & Adolescent*

- Psychology*, 35(3), 369-385. Retrieved from
https://doi.org/10.1207/s15374424jccp3503_3.
- Fabiano, G. A., Pelham, W. E., Jr., Coles, E. K., Gnagy, E. M., Chronis-Tuscano, A., & O'Connor, B. C. (2009). A meta-analysis of behavioral treatments for attention-deficit/hyperactivity disorder. *Clinical Psychology Review*, 29(2).
 doi:10.1016/j.cpr.2008.11.001
- Faraone, S. V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J. K., Ramos-Quiroga, J. A., et al. (2015). Attention-deficit/hyperactivity disorder. *Nature Reviews Disease Primers*, 1(1), 15020. Retrieved from <https://doi.org/10.1038/nrdp.2015.20>.
 doi:10.1038/nrdp.2015.20
- Faraone, S. V., & Larsson, H. (2019). Genetics of attention deficit hyperactivity disorder. *Molecular Psychiatry*, 24(4), 562-575. doi:10.1038/s41380-018-0070-0
- First, M. B. (2010). Paradigm shifts and the development of the diagnostic and statistical manual of mental disorders: Past experiences and future aspirations. *The Canadian Journal of Psychiatry*, 55(11), 692-700. Retrieved from
<https://doi.org/10.1177/070674371005501102>. doi:10.1177/070674371005501102
- Frazier, T. W., Youngstrom, E. A., & Naugle, R. I. (2007). The latent structure of attention-deficit/hyperactivity disorder in a clinic-referred sample. *Neuropsychology*, 21(1), 45-64.
 doi:10.1037/0894-4105.21.1.45
- Frick, P. J., & Nigg, J. T. (2012). Current issues in the diagnosis of attention deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder. *Annual Review of Clinical Psychology*, 8, 77-107. doi:10.1146/annurev-clinpsy-032511-143150

Garner, A. A., O'Connor, B. C., Narad, M. E., Tamm, L., Simon, J., & Epstein, J. N. (2013a).

The relationship between ADHD symptom dimensions, clinical correlates, and functional impairments. *Journal of Developmental and Behavioral Pediatrics: JDBP*, 34(7), 469-477. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/24042078>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4014057/>. doi:10.1097/DBP.0b013e3182a39890

Garner, A. A., O'Connor, B. C., Narad, M. E., Tamm, L., Simon, J., & Epstein, J. N. (2013b).

The relationship between ADHD symptom dimensions, clinical correlates, and functional impairments. *Journal of Developmental and Behavioral Pediatrics*, 34(7), 469-477. doi:10.1097/DBP.0b013e3182a39890

Gathje, R. A., Lewandowski, L. J., & Gordon, M. (2008). The role of impairment in the diagnosis of ADHD. *Journal of Attention Disorders*, 11(5), 529-537. Retrieved from <https://doi.org/10.1177/1087054707314028>. doi:10.1177/1087054707314028

Gomez, R., Vance, A., & Gomez, R. M. (2013). Maternal ratings of the ADHD symptoms:

Subtypes versus severity in clinic-referred children and adolescents. *Journal of Attention Disorders*, 20(5), 414-423. doi:10.1177/1087054713514606

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38(5), 581-586. Retrieved from

<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-7610.1997.tb01545.x>. doi:10.1111/j.1469-7610.1997.tb01545.x

Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on revised conners parent and teacher rating scales. *Journal of Abnormal Child Psychology*, 6(2), 221-236.

doi:10.1007/BF00919127

- Graziano, P. A., McNamara, J. P., Geffken, G. R., & Reid, A. (2011). Severity of children's ADHD symptoms and parenting stress: A multiple mediation model of self-regulation. *Journal of Abnormal Child Psychology, 39*(7), 1073. doi:10.1007/s10802-011-9528-0
- Greven, C. U., Buitelaar, J. K., & Salum, G. A. (2018). From positive psychology to psychopathology: The continuum of attention-deficit hyperactivity disorder. *Journal of Child Psychology and Psychiatry, 59*(3), 203-212. doi:10.1111/jcpp.12786
- Greven, C. U., Merwood, A., van der Meer, J. M. J., Haworth, C. M. A., Rommelse, N., & Buitelaar, J. K. (2016). The opposite end of the attention deficit hyperactivity disorder continuum: Genetic and environmental aetiologies of extremely low ADHD traits. *Journal of Child Psychology and Psychiatry, 57*(4), 523-531. doi:10.1111/jcpp.12475
- Groenman, A. P., Janssen, T. W. P., & Oosterlaan, J. (2017). Childhood psychiatric disorders as risk factor for subsequent substance abuse: A meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(7), 556-569. Retrieved from <https://doi.org/10.1016/j.jaac.2017.05.004>. doi:10.1016/j.jaac.2017.05.004
- Haslam, N., Williams, B., Prior, M., Haslam, R., Graetz, B., & Sawyer, M. (2006). The latent structure of attention-deficit/hyperactivity disorder: A taxometric analysis. *Australian & New Zealand Journal of Psychiatry, 40*(8), 639-647. Retrieved from <https://doi.org/10.1080/j.1440-1614.2006.01863.x>. doi:10.1080/j.1440-1614.2006.01863.x
- Hay, D. A., Bennett, K. S., Levy, F., Sergeant, J., & Swanson, J. (2007). A twin study of attention-deficit/hyperactivity disorder dimensions rated by the strengths and weaknesses of adhd-symptoms and normal-behavior (SWAN) scale. *Biological Psychiatry, 61*(5), 700-705. doi:<https://doi.org/10.1016/j.biopsych.2006.04.040>

- Hudziak, J. J., Achenbach, T. M., Althoff, R. R., & Pine, D. S. (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research, 16*(S1), S16-S23. doi:<https://doi.org/10.1002/mpr.217>
- Hurtig, T., Ebeling, H., Taanila, A., Miettunen, J., Smalley, S. L., McGough, J. J., et al. (2007). ADHD symptoms and subtypes: Relationship between childhood and adolescent symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(12), 1605-1613. Retrieved from <https://doi.org/10.1097/chi.0b013e318157517a>.
doi:10.1097/chi.0b013e318157517a
- Jarrett, M. A., Meter, A. V., Youngstrom, E. A., Hilton, D. C., & Ollendick, T. H. (2018). Evidence-based assessment of adhd in youth using a receiver operating characteristic approach. *Journal of Clinical Child & Adolescent Psychology, 47*(5), 808-820. Retrieved from <https://doi.org/10.1080/15374416.2016.1225502>.
doi:10.1080/15374416.2016.1225502
- Joober, R., & Tabbane, K. (2019). From the neo-Kraepelinian framework to the new mechanical philosophy of psychiatry: Regaining common sense. *Journal of Psychiatry & Neuroscience : JPN, 44*(1), 3-7. doi:10.1503/jpn.180240
- Kiani, B., Hadianfard, H., & Mitchell, J. T. (2019). Attention and behavioral control skills in Iranian school children. *ADHD Attention Deficit and Hyperactivity Disorders, 11*(3), 263-270. doi:10.1007/s12402-019-00289-5
- Kuriyan, A. B., Pelham, W. E., Molina, B. S. G., Waschbusch, D. A., Gnagy, E. M., Sibley, M. H., et al. (2013). Young Adult educational and vocational outcomes of children diagnosed with ADHD. *Journal of Abnormal Child Psychology, 41*(1), 27-41. Retrieved from <https://doi.org/10.1007/s10802-012-9658-z>. doi:10.1007/s10802-012-9658-z

- Lacalle, M., Ezpeleta, L., & Doménech, J. M. (2012). DSM-oriented scales of the child behavior checklist and youth self-report in clinically referred Spanish children. *The Spanish Journal of Psychology, 15*(1), 377-387. doi:10.5209/rev_SJOP.2012.v15.n1.37344
- Lai, K. Y. C., Leung, P. W. L., Luk, E. S. L., Wong, A. S. Y., Law, L. S. C., & Ho, K. K. Y. (2011). Validation of the Chinese Strengths and Weaknesses of ADHD-Symptoms and Normal-Behaviors Questionnaire in Hong Kong. *Journal of Attention Disorders, 17*(3), 194-202. doi:10.1177/1087054711430711
- Lakes, K. D., Swanson, J. M., & Riggs, M. (2011). The reliability and validity of the English and Spanish Strengths and Weaknesses of ADHD and Normal Behavior rating scales in a preschool sample: Continuum measures of hyperactivity and inattention. *Journal of Attention Disorders, 16*(6), 510-516. doi:10.1177/1087054711413550
- Lange, K. W., Reichl, S., Lange, K. M., Tucha, L., & Tucha, O. (2010). The history of attention deficit hyperactivity disorder. *Attention Deficit and Hyperactivity Disorders, 2*(4), 241-255. doi:10.1007/s12402-010-0045-8
- Marcus, D. K., & Barry, T. D. (2011). Does attention-deficit/hyperactivity disorder have a dimensional latent structure? A taxometric analysis. *Journal of Abnormal Psychology, 120*(2), 427-442. doi:10.1037/a0021405
- Murray, D. W., Arnold, L. E., Swanson, J., Wells, K., Burns, K., Jensen, P., et al. (2008). A clinical review of outcomes of the multimodal treatment study of children with attention-deficit/hyperactivity disorder (MTA). *Current Psychiatry Reports, 10*(5), 424-431. doi:10.1007/s11920-008-0068-4

- Nigg, J. (2015). ADHD: New approaches to subtyping and nosology. *The ADHD Report*, 23(2), 6-9,12. Retrieved from <https://doi.org/10.1521/adhd.2015.23.2.6>.
doi:10.1521/adhd.2015.23.2.6
- Parker, A., & Corkum, P. (2016). ADHD diagnosis: As simple as administering a questionnaire or a complex diagnostic process? *Journal of Attention Disorders*, 20(6), 478-486.
Retrieved from <https://journals.sagepub.com/doi/abs/10.1177/1087054713495736>.
doi:10.1177/1087054713495736
- Pelham, J. W. E., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, 34(3), 449-476. doi:10.1207/s15374424jccp3403_5
- Pham, A. V., & Riviere, A. (2015). Specific learning disorders and ADHD: Current issues in diagnosis across clinical and educational settings. *Current Psychiatry Reports*, 17(6), 38.
- Pliszka, S. (2007). Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(7), 894-921.
doi:<https://doi.org/10.1097/chi.0b013e318054e724>
- Pliszka, S. R. (2003). Psychiatric comorbidities in children with attention deficit hyperactivity disorder. *Pediatric Drugs*, 5(11), 741-750. Retrieved from
<https://doi.org/10.2165/00148581-200305110-00003>. doi:10.2165/00148581-200305110-00003
- Polanczyk, G. V. (2014). Dimensionality of childhood psychopathology and the challenge of integration into clinical practice. *European Child & Adolescent Psychiatry*, 23(4), 183-

185. Retrieved from <https://doi.org/10.1007/s00787-014-0539-6>. doi:10.1007/s00787-014-0539-6
- Polderman, T. J. C., Derks, E. M., Hudziak, J. J., Verhulst, F. C., Posthuma, D., & Boomsma, D. I. (2007). Across the continuum of attention skills: A twin study of the SWAN ADHD rating scale. *Journal of Child Psychology and Psychiatry*, *48*(11), 1080-1087. doi:10.1111/j.1469-7610.2007.01783.x
- Posner, J., Polanczyk, G. V., & Sonuga-Barke, E. (2020). Attention-deficit hyperactivity disorder. *The Lancet*, *395*(10222), 450-462. Retrieved from [https://doi.org/10.1016/S0140-6736\(19\)33004-1](https://doi.org/10.1016/S0140-6736(19)33004-1). doi:10.1016/S0140-6736(19)33004-1
- Raiker, J. S., Freeman, A. J., Perez-Algorta, G., Frazier, T. W., Findling, R. L., & Youngstrom, E. A. (2017). Accuracy of Achenbach Scales in the screening of attention-deficit/hyperactivity disorder in a community mental health clinic. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*(5), 401-409.
- Reale, L., Bartoli, B., Cartabia, M., Zanetti, M., Costantino, M. A., Canevini, M. P., et al. (2017). Comorbidity prevalence and treatment outcome in children and adolescents with ADHD. *European Child & Adolescent Psychiatry*, *26*(12), 1443-1457. doi:10.1007/s00787-017-1005-z
- Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*(5), 615-620. doi:10.1007/s10979-005-6832-7
- Rigoni, M., Blevins, L. Z., Rettew, D. C., & Kasehagen, L. (2020). Symptom level associations between attention-deficit hyperactivity disorder and school performance. *Clinical*

- Pediatrics*, 59(9-10), 874-884. Retrieved from
<https://doi.org/10.1177/0009922820924692>. doi:10.1177/0009922820924692
- Robaey, P., Amre, D., Schachar, R., & Simard, L. (2007). French version of the strengths and weaknesses of ADHD symptoms and normal behaviors (SWAN-F) questionnaire. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent*(2), 80-89.
- Rohde, L. A. (2008). Is there a need to reformulate attention deficit hyperactivity disorder criteria in future nosologic classifications? *Child and Adolescent Psychiatric Clinics of North America*, 17(2), 405-420. doi:<https://doi.org/10.1016/j.chc.2007.11.007>
- Salum, G. A., Sonuga-Barke, E., Sergeant, J., Vandekerckhove, J., Gadelha, A., Moriyama, T. S., et al. (2014). Mechanisms underpinning inattention and hyperactivity: neurocognitive support for ADHD dimensionality. *Psychological Medicine*, 44(15), 3189-3201.
 doi:10.1017/S0033291714000919
- Sasser, T., Schoenfelder, E. N., & Stein, M. A. (2017). Targeting functional impairments in the treatment of children and adolescents with ADHD. *CNS Drugs*, 31(2), 97-107.
- Schatz, N. K., Aloe, A. M., Fabiano, G. A., Pelham, W. E., Jr., Smyth, A., Zhao, X., et al. (2020). Psychosocial interventions for attention-deficit/hyperactivity disorder: Systematic review with evidence and gap maps. *Journal of Developmental & Behavioral Pediatrics*, 41. Retrieved from
https://journals.lww.com/jrnldb/Fulltext/2020/03001/Psychosocial_Interventions_for.6.aspx.

- Schoenfelder, E., N., & Sasser, T. (2016). Skills versus pills: Psychosocial treatments for adhd in childhood and adolescence. *Pediatric Annals*, *45*(10), e367-e372. Retrieved from <https://doi.org/10.3928/19382359-20160920-04>. doi:10.3928/19382359-20160920-04
- Schulz-Zhecheva, Y., Voelke, M., Beauducel, A., Buch, N., Fleischhaker, C., Bender, S., et al. (2017). ADHD traits in german school-aged children: Validation of the German Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN-DE) scale. *Journal of Attention Disorders*, *23*(6), 553-562. doi:10.1177/1087054716676365
- Shekim, W. O., Cantwell, D. P., Kashani, J., Beck, N., Martin, J., & Rosenberg, J. (1986). Dimensional and categorical approaches to the diagnosis of attention deficit disorder in children. *Journal of the American Academy of Child Psychiatry* *25*(5), 653-658. doi:10.1016/s0002-7138(09)60290-9
- Sibley, M. H., Pelham, W. E., Molina, B. S. G., Gnagy, E. M., Waxmonsky, J. G., Waschbusch, D. A., et al. (2012). When diagnosing ADHD in young adults emphasize informant reports, DSM items, and impairment. *Journal of Consulting and Clinical Psychology*, *80*(6), 1052-1061. doi:10.1037/a0029098
- Sibley, M. H., Swanson, J. M., Arnold, L. E., Hechtman, L. T., Owens, E. B., Stehli, A., et al. (2017). Defining ADHD symptom persistence in adulthood: Optimizing sensitivity and specificity. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *58*(6), 655-662. doi:10.1111/jcpp.12620
- Siebelink, N. M., Bögels, S. M., Speckens, A. E. M., Dammers, J. T., Wolfers, T., Buitelaar, J. K., et al. (2021). A randomised controlled trial (MindChamp) of a mindfulness-based intervention for children with ADHD and their parents. *Journal of Child Psychology and Psychiatry*, *63*(2)165-177. doi:<https://doi.org/10.1111/jcpp.13430>

- Sonuga-Barke, E. J. S. (1998). Categorical models of childhood disorder: A conceptual and empirical analysis. *Journal of Child Psychology and Psychiatry*, 39(1), 115-133.
doi:<https://doi.org/10.1111/1469-7610.00306>
- Swanson, J., Wigal, T., Lakes, K., & Volkow, N. (2012). Attention deficit hyperactivity disorder: Defining a spectrum disorder and considering neuroethical implications. *Oxford Handbook of Neuroethics*. doi:10.1093/oxfordhb/9780199570706.013.0075
- Swanson, J. M., Arnold, L. E., Molina, B. S. G., Sibley, M. H., Hechtman, L. T., Hinshaw, S. P., et al. (2017). Young adult outcomes in the follow-up of the multimodal treatment study of attention-deficit/hyperactivity disorder: Symptom persistence, source discrepancy, and height suppression. *Journal of Child Psychology and Psychiatry*, 58(6), 663-678.
Retrieved from <https://doi.org/10.1111/jcpp.12684>.
doi:<https://doi.org/10.1111/jcpp.12684>
- Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., et al. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: history of the SNAP and the SWAN rating scales. *The International Journal of Educational and Psychological Assessment*, 10(1), 51-70.
- Swanson, J. M., Wigal, T., & Lakes, K. (2009). DSM-V and the future diagnosis of attention-deficit/hyperactivity disorder. *Current Psychiatry Reports*, 11(5), 399-406.
doi:10.1007/s11920-009-0060-7
- Thapar, A. (2018). Discoveries on the genetics of adhd in the 21st century: New findings and their implications. *American Journal of Psychiatry*, 175(10), 943-950. Retrieved from <https://doi.org/10.1176/appi.ajp.2018.18040383>. doi:10.1176/appi.ajp.2018.18040383

- Valo, S., & Tannock, R. (2010). Diagnostic instability of DSM–IV ADHD subtypes: Effects of informant source, instrumentation, and methods for combining symptom reports. *Journal of Clinical Child & Adolescent Psychology, 39*(6), 749-760. Retrieved from <https://doi.org/10.1080/15374416.2010.517172>. doi:10.1080/15374416.2010.517172
- Van Os, J., Gilvarry, C., Bale, R., Van Horn, E., Tattan, T., White, I., et al. (1999). A comparison of the utility of dimensional and categorical representations of psychosis. *Psychological Medicine, 29*(3), 595-606. doi:10.1017/S0033291798008162
- Verhulst, F. C., & Van der Ende, J. (2006). *Assessment scales in child and adolescent psychiatry*: CRC Press.
- Wang, S., Zhang, D., Fang, B., Liu, X., Yan, G., Sui, G., et al. (2021). A study on resting EEG effective connectivity difference before and after neurofeedback for children with ADHD. *Neuroscience, 457*, 103-113. doi:<https://doi.org/10.1016/j.neuroscience.2020.12.038>
- Weiss, M. D., McBride, N. M., Craig, S., & Jensen, P. (2018). Conceptual review of measuring functional impairment: Findings from the Weiss functional impairment rating scale. *Evidence-based Mental Health, 21*(4), 155-164.
- Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R., et al. (2012). Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology, 121*(4), 991-1010. doi:10.1037/a0027347
- Wolraich, M. L., Hagan, J. F., Allan, C., Chan, E., Davison, D., Earls, M., et al. (2019). Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics, 144*(4), e20192528. doi:10.1542/peds.2019-2528

- Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology, 39*(2), 204-221. doi:10.1093/jpepsy/jst062
- Zachar, P., & Kendler, K. S. (2007). Psychiatric disorders: A conceptual taxonomy. *American Journal of Psychiatry, 164*(4), 557-565. Retrieved from <https://ajp.psychiatryonline.org/doi/abs/10.1176/ajp.2007.164.4.557>.
doi:10.1176/ajp.2007.164.4.557
- Zoromski, A. K., Epstein, J. N., & Ciesielski, H. A. (2021). Unique associations between specific attention-deficit hyperactivity disorder symptoms and related functional impairments. *Journal of Developmental & Behavioral Pediatrics, 42*(5).

Figure 3.1

Flow Chart for the Data Preparation and Selection of the Study Participants Based on the Study Inclusion and Exclusion Criteria

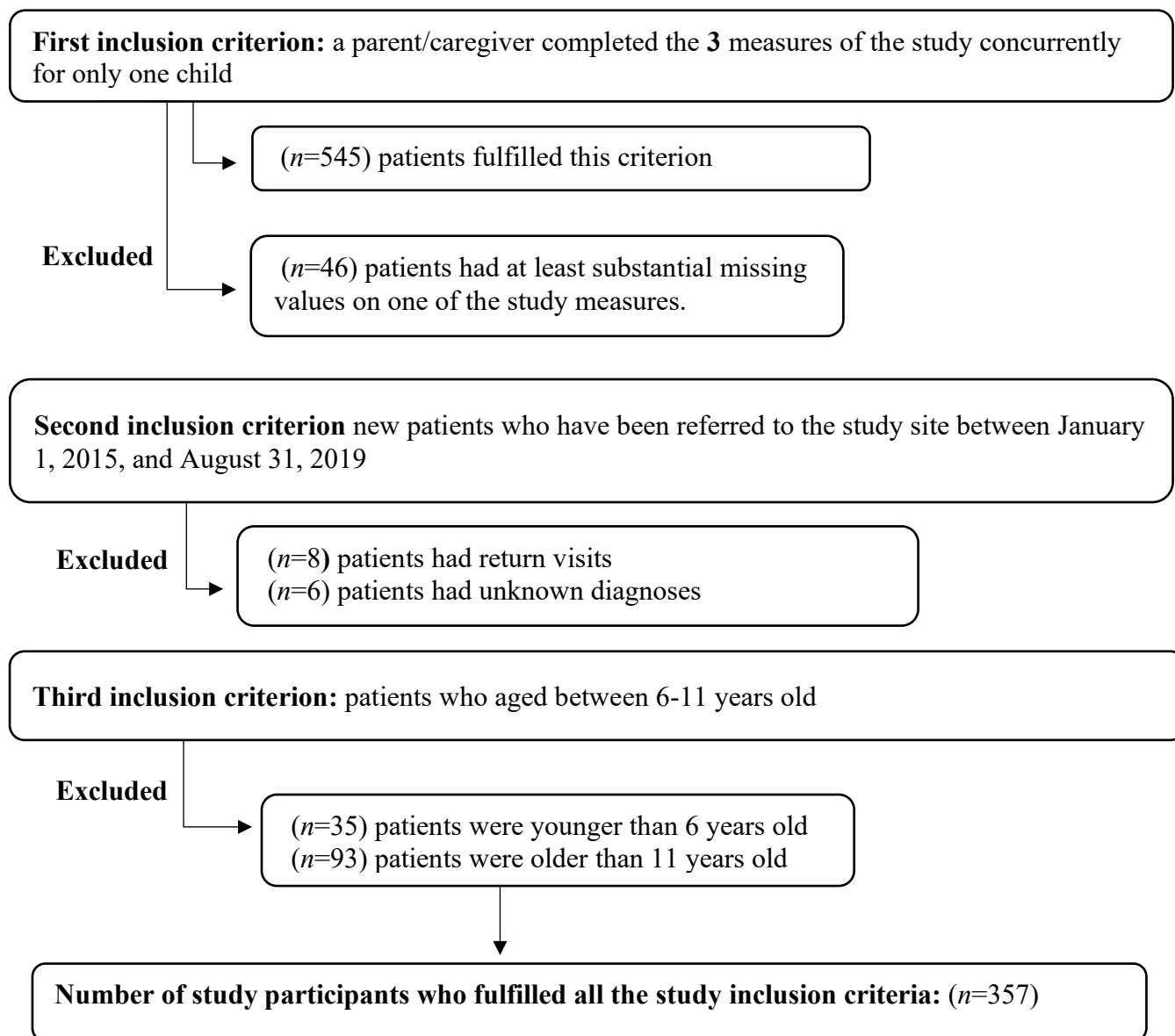


Figure 4.1.

Scatterplots between the SWAN Total and Subscales with SDQ Hyperactivity-Inattention

Subscale. Note. SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale total Score on the Inattention and Hyperactivity-Impulsivity Subscales of the SWAN scale; SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale

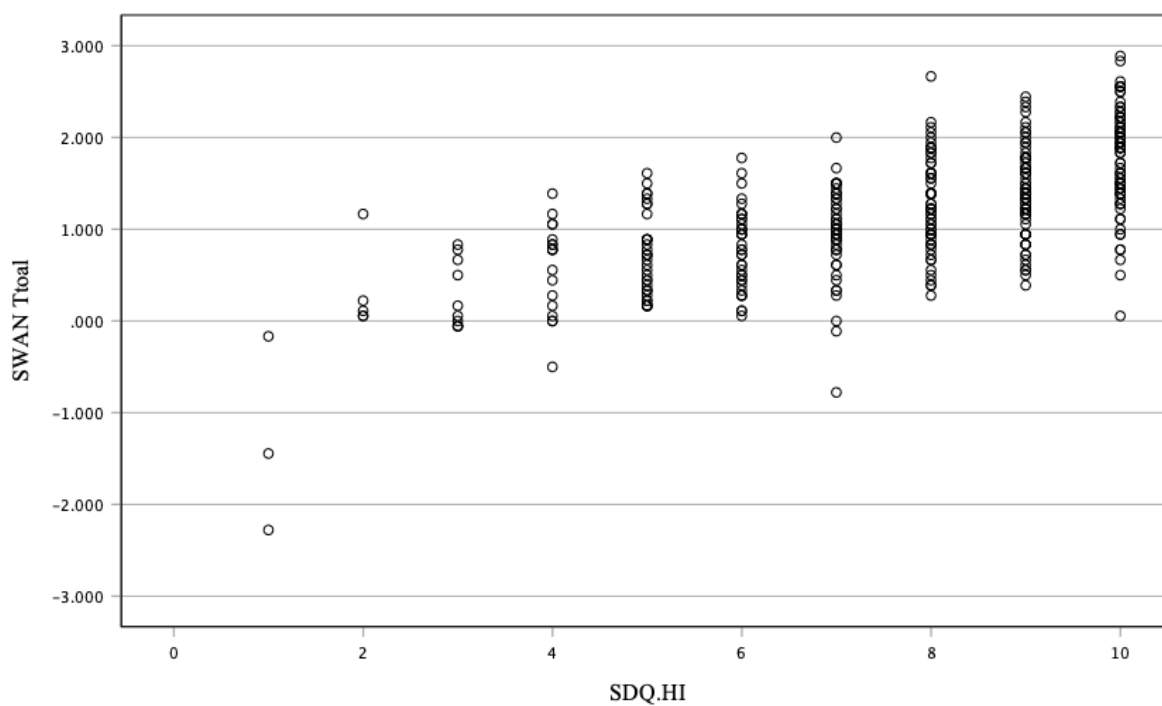
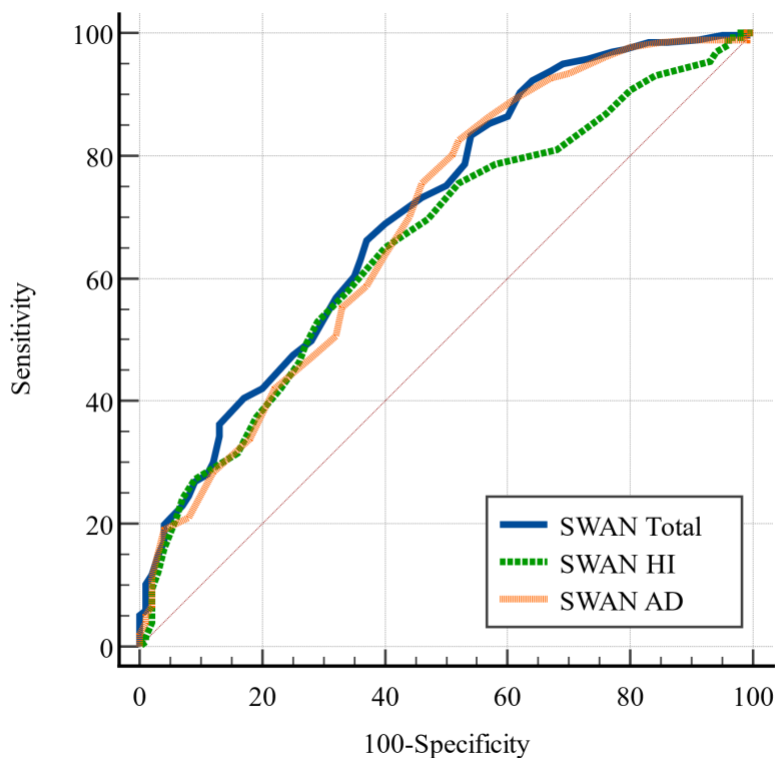


Figure 4.2.

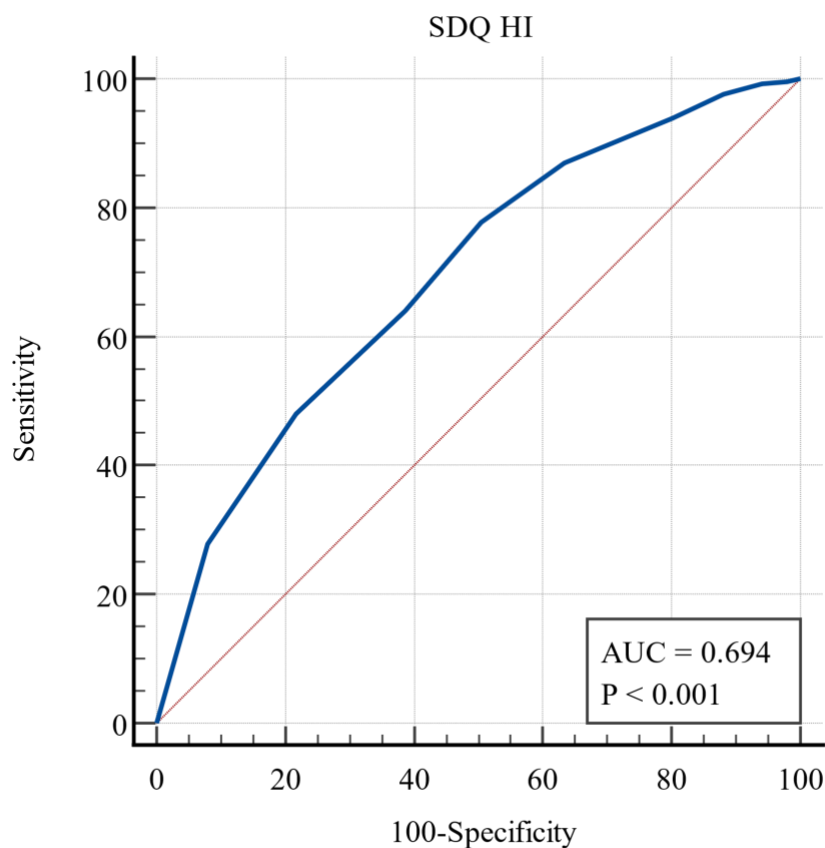
Receiver Operating Characteristic (ROC) Curves Indicating Discriminant Abilities of the SWAN Total, SWAN, HI, SWAN AD in Identifying ADHD Group (n= 257) Relative to Non-ADHD Group (n=100)



Note. SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale total score on the inattention and hyperactivity-impulsivity subscales of the SWAN scale; SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale. The diagonal line across the ROC curve graphs is a reference line and represents the characteristics of a scale which is completely useless at differentiating between those with ADHD and those without ADHD. Area Under the Curves for SWAN total, SWAN AD, and SWAN HI were (AUCs = .71, = 0.69 and 0.66, Respectively)

Figure 4.3.

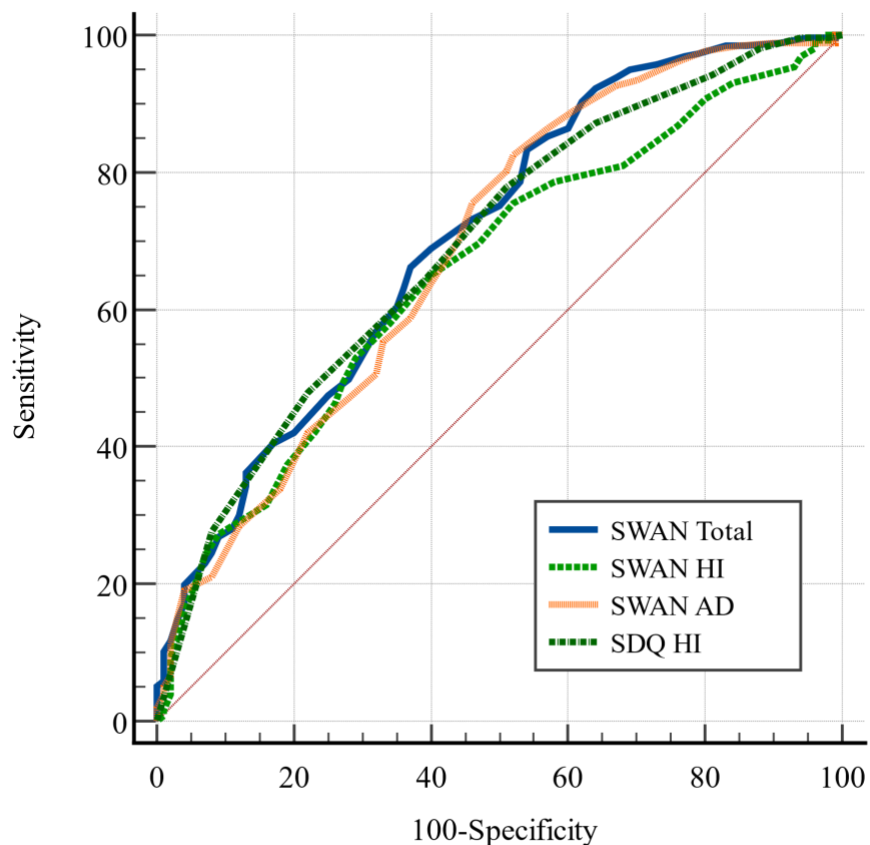
The Receiver Operating Characteristic (ROC) Curve Indicating Discriminant Abilities of The Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale in Identifying ADHD Group (n= 257) Relative to Non-ADHD Group (n=100)



Note. SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale. AUC= area under the curve. The diagonal line across the ROC curve graphs is a reference line and represents the characteristics of a scale which is completely useless at differentiating between those with ADHD and those without ADHD.

Figure 4.4.

Receiver Operating Characteristic (ROC) Curves Comparing Discriminant Abilities among SWAN Total, SWAN, HI, SWAN AD, and SDQ HI for Identifying ADHD Group (n= 257) Relative to non-ADHD Group (n=100)



Note. SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale total score on the inattention and hyperactivity-impulsivity subscales of the SWAN scale; SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale. The diagonal line across the ROC curve graphs is a reference line and represents the characteristics of a scale which is completely useless at differentiating between those with ADHD and those without ADHD.

Table 2. 1

The Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorder (DSM-5)

Symptoms of ADHD

Inattention

1. Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or with other activities
2. Often has trouble holding attention on tasks or play activities
3. Often does not seem to listen when spoken to directly
4. Often does not follow through on instructions and cannot finish schoolwork, chores, or duties in the workplace (e.g., loses focus, or is side-tracked)
5. Often has trouble organizing tasks and activities
6. Often avoids, dislikes, or is reluctant to do tasks that require mental effort over a long period of time (such as schoolwork or homework)
7. Often loses things necessary for tasks and activities (e.g., school materials, pencils, books, tools, wallets, keys, paperwork, eyeglasses, and mobile telephones)
8. Is often easily distracted
9. Is often forgetful in daily activities

Hyperactivity and impulsivity

1. Often fidgets with or taps hands or feet, or squirms in seat
2. Often leaves seat in situations when remaining seated is expected
3. Often runs about or climbs in situations where it is not appropriate (adolescents or adults might be limited to feeling restless)
4. Often unable to play or take part in leisure activities quietly
5. Is often “on the go” or acting as if “driven by a motor”
6. Often talks excessively
7. Often blurts out an answer before a question has been completed
8. Often has trouble waiting their turn
9. Often interrupts or intrudes on others (e.g., conversations or games)

Table 3.1*Demographic and Clinical Characteristics of Children in ADHD and Non-ADHD groups*

	ADHD (<i>n</i> =257)	Non-ADHD (<i>n</i> =100)	Total Sample (<i>n</i> =357)
Age (years) M(SD)	8.05(1.69)	8.17(1.67)	8.09(1.68)
Sex			
Male <i>n</i> (%)	191 (53%)	71 (20%)	262 (73%)
Female <i>n</i> (%)	66 (19%)	29 (8%)	95 (27%)
Race			
White <i>n</i> (%)	166 (46%)	67 (19%)	233 (65%)
Black <i>n</i> (%)	15 (4%)	3 (0.8%)	18 (5%)
Asian <i>n</i> (%)	27 (7%)	13 (4%)	40 (11%)
Ethnicity			
Hispanic <i>n</i> (%)	21 (6%)	7 (2%)	28 (8%)
Non-Hispanic <i>n</i> (%)	217 (60%)	85 (24%)	302 (84%)
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
SWAN Total	1.32(0.63)	0.81 (0.72)	1.18(0.70)
SWAN AD	1.46(0.75)	0.87(0.91)	1.30(0.84)
SWAN IH	1.18 (0.81)	0.74(0.80)	1.06(0.83)
SDQ HI	7.96(1.91)	6.45(2.31)	7.54(2.14)

Note. ADHD = children with ADHD diagnosis based on the clinical evaluation. Non-ADHD = children with other psychiatric or medical diagnoses based on the clinical evaluation; SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale total score

on the inattention and hyperactivity-impulsivity subscales of the SWAN scale; SWAN HI = SWAN Hyperactivity/Impulsivity subscale; SWAN AD = SWAN Attention Deficit subscale. The SWAN subscales scores range from (-3 to +3). SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale. The SDQ HI scores range from (0 to 10).

Table 4.1*Descriptive Statistics for Each of the Study Scales*

	N	Minimum	Maximum	Mean	Std. Deviation
SWAN Total	357	-2.28	2.89	1.18	0.70
SWAN HI	357	-2.22	3.00	1.06	0.83
SWAN AD	357	-2.78	3.00	1.30	0.84
SDQ/HI	357	1	10	7.54	2.14

Note. SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale total score on the inattention and hyperactivity-impulsivity subscales of the SWAN scale; SWAN HI = SWAN Hyperactivity/Impulsivity subscale; SWAN AD = SWAN Attention Deficit subscale. The SWAN subscales scores range from (-3 to +3). SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale. The SDQ HI scores range from (0 to 10). Parent = IRS item relationship with parent; Playmates = IRS item relationship with playmates; Academic = IRS item child academic progress; Family = IRS item relationship with family; Siblings = IRS item relationship with brothers and sisters; SE = Child self-esteem; and Overall = IRS item for child overall impairment. The IRS items scores range from 0 (no problem) to 6 (highly impaired).

Table 4.2

Bivariate Correlations of Convergent validity of the SWAN total and subscales with SDQ

Hyperactivity-Inattention subscale

Subscale Score	SWAN Total	SWAN HI	SWAN AD	SDQ HI
SWAN Total	—			
SWAN HI	.83**	—		
SWAN AD	.83**	.38**	—	
SDQ HI	.69**	.67**	.48**	—

Note. **: Correlation is significant at the 0.01 level (2-tailed). Pairwise was used ($n=357$)

SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN)

Scale total score on the inattention and hyperactivity-impulsivity subscales of the SWAN scale;

SWAN HI = SWAN Hyperactivity/Impulsivity subscale; SWAN AD = SWAN Attention Deficit

subscale; SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention

subscale.

Table 4.3

Bivariate Correlations of Concurrent Validity presenting correlations between IRS items, SWAN total and subscales with SDQ Hyperactivity-Inattention

IRS items	Parent	Playmates	Academic	Family	Siblings	SE	Overall
SWAN/Total	.19**	.23**	.36**	.23**	.06	.09	.31**
SWAN/HI	.22**	.30**	.19**	.23**	.07	.06	.28**
SWAN/AD	.01	.08	.41**	.14**	.04	.08	.24**
SDQ HI	.12**	.14**	.18**	.19**	.02	-.02	.23**

Note. **. Correlation is significant at the 0.01 level (2-tailed). The number of participants for the correlations ranged from ($n=357$ to 260) depending on the number of responses to the IRS items and pairwise deletion was used to maximize available sample size. SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale total score on the inattention and hyperactivity-impulsivity subscales of the SWAN scale; SWAN HI = SWAN Hyperactivity/Impulsivity subscale; SWAN AD = SWAN Attention Deficit subscale; SDQ HI = Strengths and Difficulties Questionnaire (SDQ)/ Hyperactivity-Inattention subscale. Parent = IRS item relationship with parent; Playmates = IRS item relationship with playmates; Academic = IRS item child academic progress; Family = IRS item relationship with family; Siblings = IRS item relationship with brothers and sisters; SE = Child self-esteem; and Overall= IRS item for child overall impairment. The IRS items scores range from 0 (no problem) to 6 (highly impaired).

Table 4.4

Areas Under the Curve, Sensitivities, Specificities, and Optimum Cutoff Scores for the SWAN Total, SWAN, HI, SWAN AD, and SDQ HI

	SWAN Total	SWAN AD	SWAN HI	SDQ HI
AUC	0.71	0.69	0.66	0.69
Sensitivity (%)	82.8	82.7	65.3	77.9
Specificity (%)	46.5	48.5	60.0	49.5
Optimum cutoff score	>0.72	>0.66	>0.77	>6.0

Note. Total Sample ($n=357$) ADHD n (%) / No ADHD n (%) = 257(72%) / 100 (28%). There was a significant difference in Area Under the Curves between the SWAN total and the SWAN HI ($p = .023$). No significant differences were present between the AUCs of the SWAN Total, SWAN AD, SWAN HI and SDQ HI.

SWAN Total = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) Scale total score on the inattention and hyperactivity-impulsivity subscales of the SWAN scale; SWAN HI = SWAN Hyperactivity/Impulsivity subscale; SWAN AD = SWAN Attention Deficit subscale; SDQ HI = Strengths and Difficulties Questionnaire (SDQ) / Hyperactivity-Inattention subscale. The optimum cutoff score determined by the Youden index.

Appendix A Characteristics of Studies Examined Psychometrics of the SWAN

Study	Country	Sample Characteristics				Validity					Reliabilities	
		Language	Sample Source	Gender %	Age years	Factor Structure	CV	AUC	Receiver Operative Curve Optimal Cut-off	SPE	SEN	Internal Consistency
(Burton et al., 2019)	Canada	English	Community (n=5,154)	50.8% males	6-17			CPRS-R 0.85-0.88	76-81%	77-82%	0.92-0.95	
			Clinical (n=226) Psychiatric Controls (n=36)	84.5% males				SWAN-HI: 0.60 SWAN-IA 0.72				
(Kiani et al., 2019)	Iran	Persian	Community (n= 280)	51.1% Females	6-12	CFA Two- factors SWAN-IA 0.58 -0.86% SWAN-HI 0.63 -0.81% FMM Two- classes Two-factors					0.90, 0.94	
(Gomez, Vance, & Stavropoulos, 2018)	Australia	English	Clinic-referred (n=314)	75.1% males 24.9% females	7-17							
(Schulz-Zhecheva et al., 2017)	Germany	Dutch	Community n = 343	65% male	8-18	CFA Two- factors 65%	CBCL-AP 0.87-0.91	Visual Inspection of ROC SWAN-TOT 2.86 SWAN-HI: 3.06 SWAN-IA 2.94	75%-88%	84%	0.91-0.95	27 weeks
		Online	Clinical n= 62 children with ADHD				FBB-ADHS	SWAN-TOT 2.86 SWAN-HI: 3.06				r = .77- .81
(Chan et al., 2014)	China	Chinese	Clinical n=290 children aged vs Clinical non-ADHD group	Male 140 Female 150	6-12		DISC-IV 0.92-0.95	SWAN-HI: 0.83 SWAN-IA 0.72	0.74-0.78 0.84%-0.95%	91% - 96% 79% - 96%		
(Arnett et al., 2011)	The United States	English	Colorado Twin Registry # (65%) Australian Twin Registry# (35%). n= (1,502 subjects total)	379 monozygotic and 372 same-sex dizygotic twin pairs	5-7	CFA 74%	DBRS		98%	58%	0.88	longitudinal stabilities across cross grades K, 1, 2, and 4. r=.57 to r=.75
(Lakes et al., 2011)	The United States	English and Spanish	Community (n=90)		3-5		SDQ-HI				0.95	3 months
												English SWAN rs= 0.71-0.76 Spanish SWAN rs = 0.49, - 0.61
(Lai et al., 2011)	Hong Kong	Chinese	Community (n=3,722)	51% males	6-12	Two- factors 60.1% - 81.3%		0.79-0.92	Sum of SP & SE (t-76% -89%) SWAN-TOT 0.33-0.39 SWAN-HI: 0.56-0.00 SWAN-IA 0.56-0.44	76%-89% 72% - 77%	0.90-0.95	2 to 4 weeks ICC = 84- .87

(Lakatos, Birkás, Tóth, & Gervai, 2010b)	Hungarian	Community (n= 156)	6		CBCL SDQ					0.87-0.93	
(Polderman et al., 2007)	Netherlands	Netherlands Twin Registry (NTR) (n= 561) pairs of twins	12		CBCL-AP						
(Robaey et al., 2007)	Canada	Clinically referred males (n=88)	68 (78%)	6-9	DISC-4.0	0.89					
	French		20 (22%)		Conners' Parent Conners' Teacher scores		-0.60	0.88	0.86	0.88- 0.91	

Note: All information reported in this table is about the SWAN-parent psychometric properties.

AUC/ Area under the curve :0.90-1 = **excellent** 0.80-0.90 = **good** ;0.70-.80 = **fair** ;0.60-0.70 = **poor**; 0.50-.60 = **fail** (Rice, Marnie E. & Grant T. Harris, 2005)

SPE/ Specificity

SE/Sensitivity

Studies Scales:

SWAN The Strengths and Weaknesses of ADHD-Symptoms and Normal Behavior (SWAN) Scale

SWAN/HI SWAN- hyperactivity-impulsivity subscale

SWAN/IA SWAN-inattentive subscale

SWAN/TOT/ SWAN Total Score

CPRS-R Conners' Parent Rating Scale

FBB= "Fremdbeurteilungsbogen für ADHS" (FBB-ADHS; Döpfner, Görtz-Dorten, & Lehmkuhl, 2008) is a German ADHD rating scale based on ICD-10 and *DSM-IV* diagnostic criteria.

CBCL-AP -The Child Behavior Checklist (CBCL; Achenbach, 1991) The Attention Problem subscale (CBCL-AP)

SDQ/HI Strength and Difficulties Questioner/

(DBRS) Disruptive Behavior Rating Scale

(DISC-4.0) Diagnostic Interview Schedule for Children 4.0