

**UNIVERSITY OF THE WESTERN CAPE**  
**FACULTY OF COMMUNITY AND HEALTH SCIENCES**

**Title:** A systematic review conducted on studies reporting on the instruments used in the assessment of adult ADHD.

**Student Name:** G. Robertson

**Student Number:** 3227938

**Department:** Psychology

**ORCID:** 0000-0002-5545-4857

**Supervisor:** Prof. M. Smith

**Co-Supervisor:** Ms. M. Willemsse



**Date:** 14 December 2018

Thesis submitted in fulfilment of the requirements for the M.A. Psychology (Thesis) Degree at the Department of Psychology, University of the Western Cape.

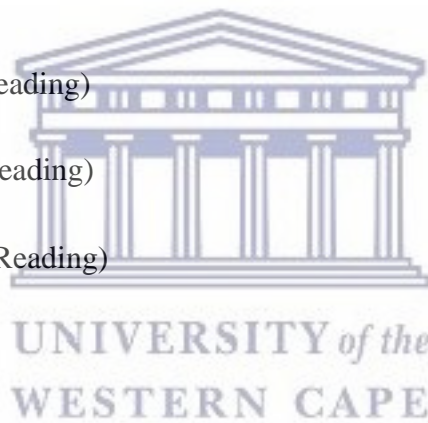
**Key Words:** adult ADHD, instruments, psychometric properties, screening, diagnosis, reliability, validity sensitivity and specificity

## Table of Contents

ACKNOWLEDGEMENTS	vi
LIST OF TABLES	vii
LIST OF FIGURES	viii
DECLARATION	ix
ABSTRACT	x
CHAPTER ONE: INTRODUCTION	1
1.1 Background	1
1.2 Problem Statement	3
1.3 Rationale	4
1.4 Aim of the study	4
1.5 Objectives of the study	4
1.6 Theoretical Framework	5
1.7 Chapter Organization	5
CHAPTER TWO: LITERATURE REVIEW	7
2.1 Diagnosis	7
2.2 Prevalence	7
2.3 Comorbidity	8
2.4 Functional Impact	9
2.5 Intervention	9



2.6 Instruments	9
<b>CHAPTER THREE: METHODOLOGY</b>	<b>12</b>
3.1 Aim	12
3.2 Objectives	12
3.3 Review Questions	12
3.4 Research Design	13
3.5 Inclusion Criteria	14
3.6 Exclusion Criteria	14
3.7 Review Process	14
3.7.1 Identification (Title Reading)	14
3.7.2 Screening (Abstract Reading)	19
3.7.3 Eligibility (Full-Text Reading)	19
3.8 Method of Review	19
3.9 Meta-Synthesis	20
<b>CHAPTER FOUR: RESULTS</b>	<b>23</b>
4.1 Process Results	23
4.2. Descriptive Meta-Synthesis	26
4.2.1 Ranking	26
4.2.2 Ranks based on subsections	27
4.2.2.1 Section A	27
4.2.2.2 Section B	33



4.3 Instruments	38
4.3.1 Types of Instruments	40
4.3.2 Format of Instruments	41
4.3.3 Versions of Instruments	43
4.3.4 Purpose of Instruments	46
4.3.5 Domains of Instruments	47
4.4 Theoretical Frameworks	52
4.5 Theoretical and Operational Definitions	52
4.6 Norms	53
4.7 Diagnostic and Statistical Manual (DSM)	53
4.8 Psychometric Properties of Instruments	57
4.8.1 Reliability	57
4.8.2 Validity	60
4.8.3 Sensitivity and Specificity	62
CHAPTER FIVE: CONCLUSION	65
5.1 Executive Summary	65
5.2 Core Findings	65
5.3 Conclusion	71
5.4 Limitations	73
5.5 Recommendation for Future Study	74
5.6 Significance of the Study	74



Reference list	76
Appendix A: Title Summary Sheet	88
Appendix B: Abstract Summary Sheet	89
Appendix C: Critical Appraisal Tool	90
Appendix D: Full Text Review Summary Sheet	95
Appendix E: Permission for SFS scoring system	96



## ACKNOWLEDGEMENTS

I would like express sincere thanks and appreciation to both the supervisors, Prof. M. Smith and Ms. M. Willemse for their expertise and emotional support. They consistently offered up time, demonstrated their commitment, provided guidance and mentorship in a highly professional manner.

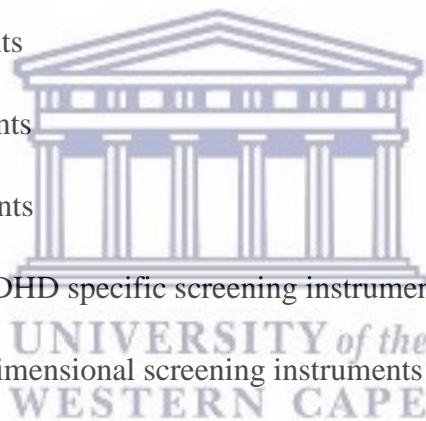
I would like to acknowledge and express thanks for the contribution made by my loving wife and children, as well as, Sabrina Maharaj for her work as a co-reviewer.

Lastly, I would like to thank the National Research Foundation (NRF) for affording me the opportunity to reach a higher level of expertise. The research presented has not been commissioned by the NRF or representing the opinions thereof.



## LIST OF TABLES

Table 3.1 List of Disciplines	16
Table 3.2 Primary Databases	17
Table 3.3 Secondary Databases	18
Table 4.1 Ranking	26
Table 4.2 Summary of Section A of appraisal tool	28
Table 4.3 Summary of Section B of appraisal tool	34
Table 4.4 Instruments identified for measuring adult ADHD	39
Table 4.5 Type of Instruments	40
Table 4.6 Format of Instruments	42
Table 4.7 Version of Instruments	44
Table 4.8 Purpose of Instruments	46
Table 4.9 Domains of adult ADHD specific screening instruments	48
Table 4.10 Domains of multidimensional screening instruments	50
Table 4.11 Domains of adult ADHD specific diagnostic instruments	51
Table 4.12 Domains of multidimensional diagnostic instruments	51
Table 4.13 Adherence to DSM	54
Table 4.14 Reliability	58
Table 4.15 Validity	60
Table 4.16 Validity in relation to study objectives	61
Table 4.17 Sensitivity and Specificity	62



## LIST OF FIGURES

Figure 3.1 PRISMA Flowchart	21
Figure 4.1 Review Process	25





## DECLARATION

I, Gershwin Robertson, hereby declare that the thesis entitled, “A systematic review conducted on studies reporting on the instruments used in the assessment of adult ADHD”, is my own work. It has not been submitted before for any degree or examination in any other university. All the sources used or quoted were acknowledged and fully referenced.



G. Robertson

14 December 2018

## ABSTRACT

The diagnosis of adult ADHD is a complex process that requires information from different sources. Instruments are used to screen or diagnose adult ADHD. The aim of the study was to identify instruments measuring adult ADHD from good quality research. This systematic review was executed following the recommended PRISMA steps. A comprehensive search was conducted across identified databases. The SFS scoring system was used to critically appraise for methodological rigour and coherence. Meta-synthesis was used to summarize extracted data from 26 articles included in the final summation. Ethics clearance was issued by the UWC Senate Research Committee. Sixteen instruments measuring adult ADHD were identified. Screening instruments measure core symptoms whereas diagnostic instruments assess all criteria. Fourteen instruments were based on DSM-IV criteria and four were based on DSM-V criteria for adult ADHD including rival explanations for the symptoms. The lack of adoption of DSM-V criteria remains a concern given criticism against DSM-IV criteria for adult ADHD. Overall instruments presented acceptable psychometric properties. However, the performance of the instruments was study dependent. A cautionary note is that these indices must be interpreted carefully. Further research must explore the reasons underlying the lack of adoption of DSM-V criteria in research, and the lack of revision of instruments measuring adult ADHD.

# CHAPTER ONE: INTRODUCTION

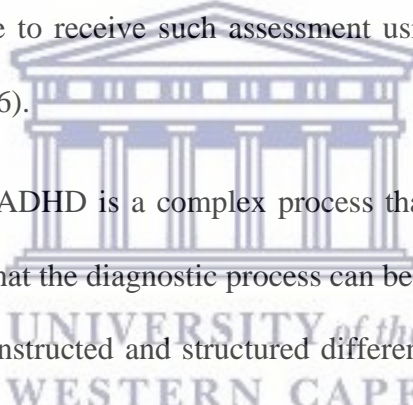
## 1.1 Background

Attention Deficit Hyperactivity Disorder (ADHD) is a childhood-onset disorder considered to be highly disruptive (Price & Raffelsbauer, 2012). The core symptomatology associated with ADHD is inattention, hyperactivity and/or impulsivity (Brod, Johnston, Able & Swindle, 2006). The ramifications of such symptoms are that they inevitably lead to difficulties that deter effective task-orientated behaviour (Wilens, Faraone & Biederman, 2004). Thus, individuals with ADHD are highly sensitive to distractions, and experience difficulty in adequately responding to demands placed on them (Sadock et al, 2015).

For years, ADHD was thought of as primarily a childhood disorder (Grinell, 2011), and consequently, was referred to as a childhood neurodevelopmental disorder (Price & Raffelsbauer, 2012). The rationale for such thought was that the symptomatology is more noticeable in childhood (Van der Westhuizen, 2010). The assumption was that ADHD subsides in adolescents, and thus fails to continue into adulthood. Grinell (2011) argued that this is however not the case. There is an increasing recognition that ADHD often continues into adulthood (Grinell, 2011; McGough & Barkley, 2004).

The recognition of ADHD continuing into adulthood has subsequently seen great interests in research on adult ADHD, particularly the diagnosis of adult ADHD (Atwoli, Owiti, Manguro & Ndambuki, 2011; Grinell, 2011; McGough & Barkley 2004). These studies have been imperative in highlighting and making more information available on this phenomenon. Perhaps most crucial, is the identification of the symptom changes that has become apparent in the development of ADHD throughout childhood, adolescent and adulthood. This changing symptomatology implies that as one ages, the symptoms of ADHD inevitably change, and as a

consequence, ADHD manifests slightly differently in the various developmental stages (Weiss & Weiss, 2004). According to the Diagnostic and Statistical Manual of Mental Disorders (DSM 5) the main manifestation in preschool years is hyperactivity (American Psychiatric Association, 2013). This differs to elementary school years where inattention is more prominent; adolescents are characterized by fidgetiness, an inner feeling of jitteriness, or impatience. Adulthood presents with symptoms of inattention and restlessness (American Psychiatric Association, 2013). Identifying how the symptoms manifest differently has led to the formation of separate diagnostic tools, some relevant to childhood symptoms, and others specific for adolescents and adult symptoms (Epstein & Kollins, 2006). Accordingly, individuals who have not been diagnosed with ADHD in childhood, and present a need to be assessed in adulthood are able to receive such assessment using adult ADHD measurement scales (Epstein & Kollins, 2006).



Diagnosing adults with ADHD is a complex process that requires retrospective recall. Rösler et al (2006) identified that the diagnostic process can be informed by a host of different instruments that have been constructed and structured differently, therefore operating within unique properties. Clinicians are largely dependent on the psychometric properties of the scales to make an informed diagnosis (McCann & Roy-Byrne, 2004). Thus, the importance of conducting inquiries regarding the psychometric properties of these adult scales becomes apparent. Knowing which scales are most effective would greatly assist clinicians in their diagnostic process.

Many studies have provided important information regarding the validity and reliability of adult ADHD scales (e.g. Kessler et al, 2007; Kooij et al, 2013; Rösler et al, 2006; Takeda et al, 2015; Yeh, Gau, Kessler & Wu, 2008). However, the majority of such studies focused solely on individual scales. There have been attempts at filtration in order to consolidate the literature

on instruments measuring adult ADHD. For example, Taylor, Deb and Unwin (2011) conducted a systematic review reporting on the psychometric properties of scales measuring adult ADHD.

## **1.2 Problem Statement**

There have been numerous studies conducted using scales to measure adult ADHD. These primary studies report specifically on certain assessment scales used (e.g. Sanchez-Gacia et al, 2015; Marchant, Reimherr, Robison, Robison & Wender, 2013; Spencer et al, 2009; Yeh, Gau, Kessler & Wu, 2008), whilst other studies report specifically on the construction of these scales (e.g. Eardt, Epstein, Conners, Parker & Sitarenois, 1999; Mehringer et al, 2002; Watson & Liljequist, 2015). There has been an attempt at filtration of the literature reporting on instruments measuring adult ADHD. However, the only other systematic review was published in 2011 and reviewed publications retrieved from four medical journal databases, each with distinct search periods; 1) Medline (1950 – June 2010); 2) Cinahl (1981 – June 2010); 3) Embase (1980 – June 2010); and 4) Psych Info (1967 – June 2010) (Taylor et al, 2011). This study was further limited, as their appraisal of the included primary research articles indicated that many lacked good quality (Taylor et al, 2011). The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) was used to judge the quality of the studies (Taylor et al, 2011). However, none of the studies were excluded on the basis of study design or the quality of study since only 35 studies were retrieved. They identified fourteen different instruments and found that the Conners' Adult ADHD Rating scale and the Wender Utah Rating Scale (short version) had more robust psychometric statistics (Taylor et al, 2011). Nonetheless, it was further recommended that a meta-analysis be performed to support their findings (Taylor et al, 2011). Thus, a need for filtered information exists for the period 2010 – 2016, searching other databases, selecting only primary research that is of good quality and with a more nuanced focus including the sensitivity and specificity of these instruments. The aim is to produce filtered information based on evidence from good quality

research. Information that serves to validate the employment of certain scales over the use of others is essential, as cautions are raised against the utilisation of any adult ADHD measurement scale.

### **1.3 Rationale**

The diagnosis of adult ADHD has been identified as a complex task (Rösler et al, 2006). Taylor et al (2011) provided evidence that further research be conducted on the validity of adult ADHD rating scales. The intention is to present adequate scales for diagnostic accuracy—scales that epitomize sensitivity when assessing adults with ADHD, as well as specificity in cases of non-ADHD adults. The identification of such scales is dependent on filtered information assessing the quality of the psychometric properties of these instruments.

### **1.4 Aim of the study**

The aim of this study was to consolidate the literature reporting on instruments used in the assessment of adult ADHD.

### **1.5 Objectives of the study**

- To identify instruments measuring adult ADHD from good quality research
- To identify domains included in the measurement of the construct ADHD
- To identify the theoretical frameworks adopted in studies using instruments measuring adult ADHD
- To identify and summarize the theoretical and operational definitions provided for the respective measures
- To identify and summarize the reference or norm groups for the respective measures
- To identify and summarize how the instruments reflect revisions in the diagnostic criteria in DSM-V



- To identify and summarize the psychometric properties of the respective measures including a) Reliability, b) Validity, c) Sensitivity, and d) Specificity

## **1.6 Theoretical framework**

It is a requirement at UWC that full thesis masters have theoretical frameworks. This is predicated on the assumption that full masters studies are reactive ones in which data is collected from live participants. The topic of the present study overlaps two main theoretical domains, 1) Test construction, that includes a limited range of theory with greater emphasis on the operational steps, and 2) Psychopathology, that includes developmental theory and dimensional conceptualization for syndromes. In both instances, the focus is on the development of symptoms and how these symptoms hang together to form a syndrome, as well as how to assess whether that syndrome reaches disorder level.

The present study drew on both the theoretical underpinnings of test construction and developmental psychopathology. Given that the present study also adopted a secondary research design, it was more feasible to integrate these elements into the procedures of the thesis than to explicitly adopt it as a theoretical framework. It would be problematic to force the data collection and interpretation of results into these theoretical frameworks, as the primary studies included in the final summation all adopted their own theoretical perspectives. Thus, the work will be evaluated for methodological rigour, coherence and interpreted from their respective theoretical frameworks. In this way the deviation from the requirement to include a theoretical framework is reported here as an explicit methodological decision in the interest of maintaining the integrity of the source data.

## **1.7 Chapter organization**

The thesis is presented in five chapters. The first chapter provides an overview of the study and the problem statement that the research attempted to address. The second chapter

provides a brief literature overview. This chapter attempted to provide an academic rationale for the present study. The third chapter reports on the methodological choices made and the process of conducting the study. This chapter attempts to provide enough information to make evaluation of the methodological rigour and coherence possible, as well as to make replication possible. The fourth chapter presents the results of the study. The fifth chapter presents the discussion and conclusion of the study findings.





## CHAPTER TWO: LITERATURE REVIEW

The literature reporting on adult ADHD includes the following: a) diagnosis, b) prevalence, c) comorbidity, d) functional impact, e) intervention and f) measurement of ADHD. Below follows a brief overview of the literature on adult ADHD.

### 2.1 Diagnosis

Attention-Deficit/Hyperactivity Disorder (ADHD) in adults has been increasingly recognised by researchers and clinicians (Liebenberg, 2016; Ramsay 2017). Nevertheless, there are many uncertainties regarding the process and accuracy of the diagnosing of adult ADHD relating to the unreliability of self-reflection and retrospective recall of ADHD symptoms (Kooij et al, 2008; Taylor et al, 2011), the high comorbid presence of other disorders co-occurring with ADHD (e.g. Liebenberg, 2016), and questions on the lack of an age-appropriate criteria (e.g. Asherson et al, 2012). Five changes have been made in the DSM-V to provide more accurate criteria for adult ADHD (APA, 2013). Prosser and Reid (2013) concluded that there remains a great potential for misdiagnosis despite the more acceptable criteria for adult ADHD in the DSM-V. Liebenberg (2016) concluded that DSM-V criteria remain insufficient to provide a complete account of adult ADHD. One of the main concerns of the DSM is that it is considered to be more appropriate for child ADHD and thus does not adequately identify ADHD presentations in adults (Walls, Wallace, Brothers & Berry, 2017). The DSM-V has addressed some of these issues, however, instruments have not reflected these changes (Walls et al, 2017). It is recommended that the diagnosis of the syndrome remain a focus of further research.

### 2.2 Prevalence

Prevalence of adult ADHD is not as well documented as childhood ADHD (Atwoli, Owiti, Manguru & Ndambuki, 2010; Walls et al, 2017). Initially, the prevalence was expected

to be lower as ADHD was thought to remit by young adulthood (Hesson & Fowler, 2012). Atwoli et al (2010) attributed lower prevalence rates to inappropriate diagnostic criteria that were not developed for adults resulting in mis- and underdiagnosis. The accepted estimation of the worldwide prevalence for adult ADHD ranges between 3% and 5% (Asherson et al, 2012; Atwoli et al, 2010; Gjervan, Torgersen, Nordahl & Rasmussen, 2012; Sprafkin, Gadow, Weiss, Schneider & Nolan, 2007; Van Schalkwyk & Schronen, 2011).

Recent studies indicate that adult ADHD is more prevalent with estimates ranging from 30% to 78% of the cases including both new diagnoses and childhood diagnoses that persist into adulthood (Asherson et al, 2012; Gjervan et al, 2012; Van Schalkwyk & Schronen, 2011). The increase in new adult diagnoses was attributed to the changes of diagnostic criteria in the DSM-V that account better for the experience of adult ADHD (Hesson & Fowler, 2012). Thus, the accuracy of diagnosis and assessment remain a focus of further research.

### **2.3 Comorbidity**

Research indicates that as much as 90% of patients with ADHD presents with one or more comorbid disorder (APA, 2013; Giacobini, Medin, Ahnemark, Russo & Carlquist, 2014; Nelson & Gregg, 2012). The most common psychiatric disorders that co-occur with adults meeting the criteria for ADHD are anxiety disorders, mood disorders, substance abuse disorders, antisocial personality disorders and developmental disorders such as learning disabilities (Kooij et al, 2012; Gjervan, 2012). Comorbidity serves as one of the key challenges to diagnosing adult ADHD (Giacobini et al, 2014; Weyandt & DuPaul, 2012). Nelson and Gregg (2012) recommended continued examination of the extent to which diagnostic processes account for comorbidity in adult ADHD.

## **2.4 Functional Impact**

Adult ADHD is associated with poor functional outcomes, educational difficulties, lower occupational accomplishment, unemployment, and greater risk of workplace injury (Alderson, Hudec, Patros & Kasper, 2013; Barkley & Murphy, 2010; Gjervan et al, 2012). Adult ADHD was associated with increased marital and family problems, lower perceived social competence and increased criminal behaviour, as well as increased intimate partner violence (Hesson & Fowler, 2012; Sibley et al, 2012). Ramsay (2016) linked adult ADHD to low self-esteem, low life satisfaction and poor physical health. Functional impairments can affect the diagnosis by overshadowing the primary symptoms of ADHD (Sibley et al, 2012).

## **2.5 Intervention**

Literature states that pharmacological forms of intervention are considered to be at the front line of treating ADHD with differential views on stimulant and non-stimulant treatment (Alder, Solanto, Escobar, Lipsuis & Upadhyaya, 2016; Prevatt & Young, 2014; Sadock et al, 2015). Medication effectively produces symptom improvements, but rarely functional improvements (Halperin, Bédard & Lichtin, 2012). Pharmacological treatment is recommended with adjunctive interventions targeting different domains such as academic, workplace and interpersonal to name a few (Ramsay, 2016).

## **2.6 Instruments**

Epstein and Kollins (2006) concluded that the development of assessment tools for diagnosing adult ADHD has lagged behind childhood ADHD, and has been adapted from the best practice guidelines for assessing childhood ADHD. Literature explicitly stated that instruments assessing adult ADHD must also assess for symptoms that were present since or during childhood (Grinell, 2011; Kooij et al, 2008; McCann, Scheele, Ward & Roy-Byrne

2000; McGough & Barkely, 2004; Taylor et al, 2011). Retrospective scales assess historical functioning and symptom presentation (Marchant et al, 2013).

Instruments can either take on the form of self-report or clinician-administered scales (Epstein & Kollins, 2006; Marchant et al, 2013). Self-reports are completed by the index patient or an informant (e.g. a parent or spouse) who provides a description of the symptoms during childhood (De-Quiros & Kinsbourn, 1998). Kooij et al (2008) highlighted that adults with ADHD often disagree with the report or description provided by the informant ostensibly due to a reported tendency of adults with ADHD to underreport problems related to inattention. Informants can also be unaware of the internal problems faced by adults with ADHD (Taylor et al, 2011).

Clinician rated scales take the form of interviews that allow a clinician to attain relevant information for diagnosis (Marchant et al, 2013). The ability of the clinician to complete the scale and make an informed diagnosis is dependent on the descriptions of an informant or the self-report of the adult being diagnosed (Rösler et al, 2006). Taylor et al (2011) identified the robustness of scales as a focus of further research in adult ADHD.

As mentioned before, instruments have been used in the clinical and research arena to measure adult ADHD (e.g. Sanchez-Gacia et al, 2015; Yeh, Gau, Kessler & Wu, 2008). Taylor et al (2011) reported on a systematic review of instruments measuring ADHD. These authors concluded that the need for filtered information continued and recommended an expansion of the databases, meta-analysis and more rigorous investigation of psychometric properties as foci of further research. McCann and Roy-Byrne (2004) identified the importance of the sensitivity (correctly identifying adults as having ADHD) and specificity (correctly identifying adults as Non-ADHD) of scales in making accurate diagnoses. Furthermore, the review only covered until 2010, did not distinguish between versions of the DSM and does not reflect the changes

in diagnostic criteria in the current revision of the DSM. Thus, the present study aimed to filter information for the period 2012 – 2016 based on evidence from good quality research using an expanded search strategy and a more nuanced focus on psychometric properties, including the sensitivity and specificity of these instruments. In addition, the present study aimed to add critical appraisal for methodological rigour and coherence. In contrast to Taylor (2011) who omitted the critical appraisal due to low number of articles identified.



## CHAPTER THREE: METHODOLOGY

### 3.1 Aim

The aim of this study was to consolidate the literature reporting on instruments used in the assessment of adult ADHD.

### 3.2 Objectives

- To identify instruments measuring adult ADHD from good quality research
- To identify domains included in the measurement of the construct ADHD
- To identify the theoretical frameworks adopted in studies using instruments measuring adult ADHD
- To identify and summarize the theoretical and operational definitions provided for the respective measures
- To identify and summarize the reference or norm groups for the respective measures
- To identify and summarize how the instruments reflect revisions in the diagnostic criteria in DSM-V
- To identify and summarize the psychometric properties of the respective measures including a) Reliability, b) Validity, c) Sensitivity and d) Specificity

### 3.3 Review questions

- What instruments measuring adult ADHD can be identified from good quality research?
- What domains are included in the measurement of the construct ADHD?
- What theoretical frameworks were adopted in studies using instruments measuring adult ADHD?
- What theoretical and operational definitions are used in measures of adult ADHD?
- What reference or norm groups are used in measures of adult ADHD?

- How do the respective measures reflect revisions in the diagnostic criteria in DSM-V?
- What are the psychometric properties of the measures of adult ADHD?
  - What are the reliability estimates?
  - Which validity indicators are reported?
  - What are the sensitivity estimates of the measures?
  - What are the specificity estimates of the measures?

### **3.4 Research Design**

This study utilized systematic review methodology. This design was deemed most appropriate, as it enabled an evaluation and interpretation of relevant literature (Higgins & Green, 2006). This was done in such a way that researchers were able to produce a comprehensive and unbiased account on the topic of inquiry (Teing, 2007). In other words, the researcher identified literature on a specific topic e.g. the psychometric properties of scales employed in the assessments of adult ADHD. This literature was then evaluated for methodological rigour and coherence relative to a threshold criterion, and the information derived as a result of such appraisal is recognised as good quality evidence (Higgins & Green, 2006). In fact, a systematic review is considered to be the highest form of evidence, as it is encapsulated as an overview of primary research (Teing, 2007). The systematic review methodology contains a statement of objectives, materials and methods, which have been produced through unambiguous, clear and reproducible methods (Teing, 2007). This was considered a fitting methodology for this study in that it enabled the filtered information of the relevant studies reporting on the psychometric properties of scales used in the diagnosis of adult ADHD. In this way, the production of filtered information or evidence has provided validation for the utilization of certain scales.

### **3.5 Inclusion Criteria**

*Type of Participants:* The review only included studies with adult participants i.e. over the age of 18.

*Time period of review:* The proposed study included literature from January 2012 to August 2016.

*Types of Studies:* The systematic review included studies that incorporated the use of a scale that measures adult ADHD. Design or type of research study was not used as an exclusionary criterion.

*Text selection:* Only full text articles housed in the identified databases were included in this systematic review. Only studies published in English were considered.

### **3.6 Exclusion Criteria**

Articles that were not published in English, e.g. foreign languages and other local languages, were excluded. Studies that required payment, or where full text was not available, were excluded from this systematic review. If the studies were not peer reviewed, they were excluded. If the studies were published outside of the time frame, they were excluded.

### **3.7 Review Process**

This study used a 3-step review process namely: 1) Identification (title reading), 2) Screening (abstract reading) and 3) Eligibility (full-text reading). Each step in this process included operational steps. Below is a description of the steps in the process.

#### **3.7.1 Identification (Title Reading)**

In this phase of the process the titles for inclusion were identified. It included three operational steps namely: a) Keyword Identification, b) Database Selection, and c) Database Search.



**A. Keyword Identification:** A list of keywords was refined from a provisional list of keywords identified from literature related to this particular study. The provisional list of keywords was: adult ADHD, scales, measurements, instrument, and psychometric properties. An exploratory search was conducted to test the provisional list of keywords on Google Scholar and PsycArticles. The effectiveness of this provisional list was determined by the number of relevant articles yielded in this initial search. Related search terms were then identified, tested and refined until a final list of search terms was selected. The final list of keywords was: instruments, tools, scales, psychometric properties and adult ADHD.

The keywords were then combined using Boolean functions such as: AND, OR and NOT to create phrases or strings. These Boolean phrases enable the researcher to narrow or broaden their search and were therefore thought to produce search results that were more focused and relevant consistent with the recommendation of Brusilovsky, Ahn, and Rasmussen (2010). These Boolean strings were adjusted and tested on Google Scholar and PsycArticles to determine which phrases would yield optimal results. The final list of Boolean strings decided upon were as follows:

Boolean phrase 1: psychometric properties OR (instruments or scales) “adult ADHD”

Boolean phrase 2: psychometric properties OR (instruments or tools) “adult ADHD”

Boolean phrase 3: psychometric properties OR (tools or scales) “adult ADHD”

**B. Database Selection:** The final list of Boolean phrases was used to conduct a comprehensive search across various databases on the UWC Website. These databases are arranged according to a number of disciplines on the UWC library website. ADHD is a well-known subject across specific disciplines. A list of the disciplines linked to ADHD are reflected in Table 3. 1 below.

**Table 3.1 List of Disciplines**

<b>Discipline</b>
Nursing
Occupational Therapy
Psychology
School of Pharmacy
School of Natural Medicine
School of Public Health
Social work

Table 3.1 provided the list of disciplines on the UWC website that was linked to ADHD, as well as their corresponding primary and secondary databases. The databases were compared across disciplines. The databases that occurred more frequently across the above-mentioned disciplines were categorized as primary databases. Table 3.2 below provides the final list of primary databases that were searched in this systematic review. It includes the Boolean phrases that were searched with the respective databases.

**Table 3.2: List of Primary Databases**

<b>Code</b>	<b>Primary</b>	<b>Keywords utilized in search</b>
<b>A</b>	BioMed Central	psychometric properties OR scales AND adult ADHD; diagnosing or assessment AND adult ADHD
<b>B</b>	Cambridge Journals Online	psychometric properties OR scales AND adult ADHD; diagnosing or assessment AND adult ADHD
<b>C</b>	Cochrane Library	psychometric properties OR scales AND adult ADHD
<b>D</b>	Credo Reference	instruments OR scales AND adult ADHD
<b>E</b>	Ebscohost	psychometric properties OR scales AND adult ADHD; diagnosing or assessment AND adult ADHD
<b>F</b>	Emerald eJournals Premier	instruments OR scales AND adult ADHD; psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD
<b>G</b>	Google scholar	psychometric properties OR scales AND "adult ADHD"; diagnosing AND "adult ADHD"
<b>H</b>	Jstor	psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD
<b>I</b>	Oxford Journals online	scales and adult ADHD; ADHD in adults
<b>J</b>	Wiley Online	psychometric properties OR scales AND adult ADHD; diagnosis and adult ADHD
<b>K</b>	Sabinet Reference	psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD;
<b>L</b>	SAGE Journals Online	psychometric properties OR scales AND adult ADHD; scales and adult ADHD; Diagnosis or screening AND adult ADHD; adult ADHD or ADHD in adults
<b>M</b>	ScienceDirect	psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD
<b>N</b>	Scopus	psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD
<b>O</b>	SpringLink	psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD

Databases that only occurred across some disciplines were categorized as secondary databases.

Table 3.3 summarizes the secondary databases and the Boolean phrases used to search.

**Table 3.3: Secondary databases**

Code	Database	Boolean phrases
P	Access Pharmacy (ONLY BOOKS)	psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD; ADHD in adults; adult ADHD (ONLY BOOKS)
Q	Agricola	psychometric properties OR scales AND adult ADHD; diagnosis or screening AND adult ADHD; ADHD in adults; adult ADHD
R	Annual Reviews	diagnosis OR screening AND ADHD in adults; scales AND ADHD in adults; psychometric properties OR scales AND adult ADHD
S	Article first (SABINET)	psychometric properties OR scales AND adult ADHD; diagnosis and ADHD; psychometric properties and ADHD; ADHD
T	Biological Abstracts	psychometric properties OR scales AND adult ADHD; diagnosis and ADHD; scales and ADHD; adult ADHD; ADHD in adults; ADHD
U	Current & Complete Research (SABINET)	Subsumed in Sabinet
V	Medicine Complete (subscription required)	psychometric properties OR scales AND adult ADHD; diagnosis and ADHD; scales and ADHD; adult ADHD; ADHD in adults; ADHD
W	South African Portals	psychometric properties OR scales AND adult ADHD; Diagnosis or screening AND adult ADHD
X	The African Journal Archive	psychometric properties OR scales AND adult ADHD; diagnosing or assessment AND adult ADHD

**C. Database Search:** The Boolean strings were used to conduct a comprehensive search across the primary and secondary databases included in the composite list in the previous step. All duplicate titles were removed. The titles of all literature identified, with specific information required for locating the texts, were recorded on a title summary sheet. The titles were then reviewed for suitability and further inclusion to this study. The reviewers conducted this step independently. Once they completed, they conferred in order to make final decisions about titles. The decisions were also recorded on the title summary sheet (Appendix A).

### **3.7.2 Screening (Abstract Reading)**

The second step in the review process was the screening of the abstracts of articles included in the title review. The abstracts were evaluated relative to the inclusion and exclusion criteria. The outcome of the screening was then recorded on the abstract summary sheet (Appendix B). Abstracts that met the criteria were forwarded to the next step in the process and those that did not meet these criteria were excluded.

### **3.7.3 Eligibility (Full-Text Reading)**

Full texts of the abstracts included from the previous step will be retrieved. This step will entail the evaluation of the texts using a critical appraisal tool. The SFS scoring system developed by Smith, Franciscus, Swartbooi, Munnik and Jacobs (2015) was used. For the purpose of this study, version D, associated with psychometric properties, was deemed most appropriate. Version D of the tool comprises two main sections namely, methodological rigour and instrumentation (Appendix C). The overarching goal of the critical appraisal tool is to evaluate the methodologies used in studies and to award scores based on specific criteria. Each article obtains a total score that is expressed as a percentage. Total scores can be categorized as weak (0-40%), moderate (41-60%), strong (61-80%) or excellent (81-100%) (Smith, 2015). A threshold score of 60% (falling within or above the strong category) was set. In other words, all articles that obtained a total score of 61% and higher was eligible for inclusion in the final review. Version D has been piloted in more than 20 studies with excellent functioning (Munnik & Smith, 2015). The outcome of this step was recorded on a full text review summary sheet (Appendix D)

### **3.8 Method of Review**

The literature retrieved from the databases were evaluated simultaneously by two reviewers. The reviewers were responsible for the evaluation and documentation of their

findings independently. This was consistent with the recommendation of Godfrey and Harrison (2012) that systematic reviews be conducted by at least two reviewers, and contributed to the rigour of the methodology. Each level of assessment was contrasted by the evaluators, and inconsistencies between reviewers were resolved through discussion. This discussion was intended to elucidate the inconsistencies. The decision was made that the supervisor make the final decision when the lack of agreement persisted. However, there were no such instances.

### **3.9 Meta-synthesis**

According to Walsh and Downe (2005), a meta-synthesis involves systematically integrating the findings yielded from inter-related individual studies. A descriptive meta-synthesis was used in the present study. Sandelowski, Docherty and Emden (1997) stated that a descriptive meta-synthesis integrates findings to generate a broad description of the research phenomenon. The descriptive meta-synthesis consisted of three components, namely the process results, rankings based on methodological rigour and instrumentation, and the synthesis. Process results entailed reporting the findings at each level of the review and related operational steps were reported. The PRISMA flow chart was adapted to illustrate the findings throughout the review process (Figure 1).

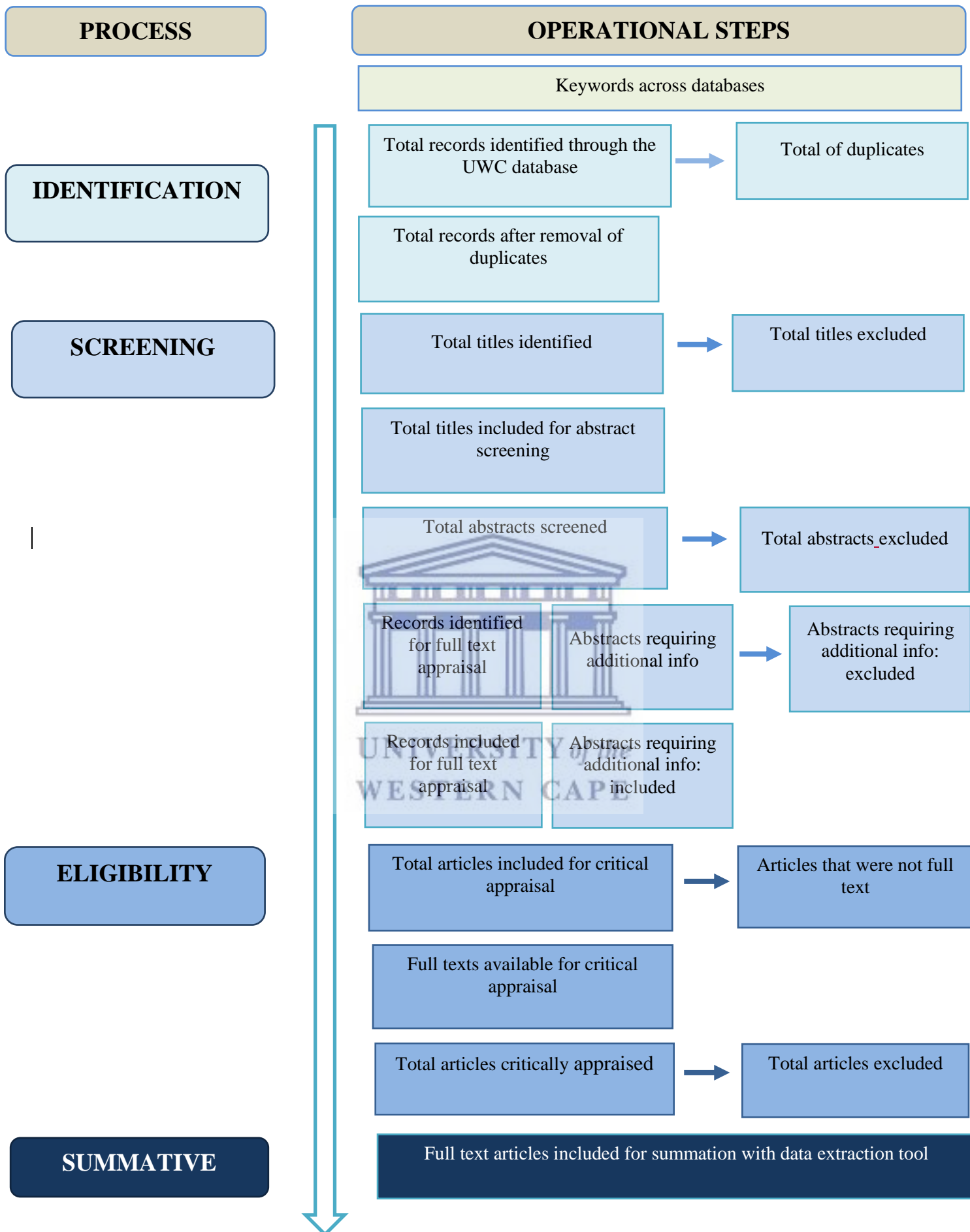


Figure 3.1: PRISMA Flowchart

The PRISMA review flow chart, developed by Moher et al (2010), refers to the Preferred Reporting Items for Systematic reviews and Meta-Analyses, and is a statement aimed toward improving the way in which systematic reviews and meta-syntheses are reported. Ranking entailed a ranked listing of articles based on total and section scores obtained on the critical appraisal tool. This enabled a nuanced engagement with the rankings obtained for methodological rigour. Synthesis entailed a summary of data extracted from eligible articles included in the final review. Data was extracted to facilitate answering the review questions.

#### **4. Ethics**

Ethics clearance and project registration was granted by the UWC Senate Research Ethics Committee (2016/8/1). The researcher must be a registered student at the institution in order to lawfully gain access to the university library and resources for data collection. Published articles retrieved are available in the public domain. No additional ethics requirements with regards to access and confidentiality. Permission to use the SFS scoring system was granted by the authors (Appendix E). The researcher submitted all revisions to the author for final approval. The present study was funded by the National Research Foundation (NRF), but does not express the opinions or views of the NRF.



## CHAPTER FOUR: RESULTS

This chapter reports on the results of this study. It consists of two main sections, namely; Process Results and Descriptive Meta-Synthesis. Process Results entailed reports on the findings at each level of the review and the operational steps undertaken. The Descriptive Meta-Synthesis includes a presentation of the ranking of all the articles based on appraised scores. In addition, it includes the summary of the data extracted from articles included in the final summation.

### 4.1 Process Results

#### Step 1: Identification

The comprehensive search across databases yielded 7519 potential articles. A total of 1970 duplicates were identified and removed, leaving 5549 potential titles. A total of 5437 titles were excluded based on their perceived lack of relevance to the review question. Thus a total of 112 titles were identified as appropriate for this review and were included in the abstract review.

#### Step 2: Screening (Abstract reading)

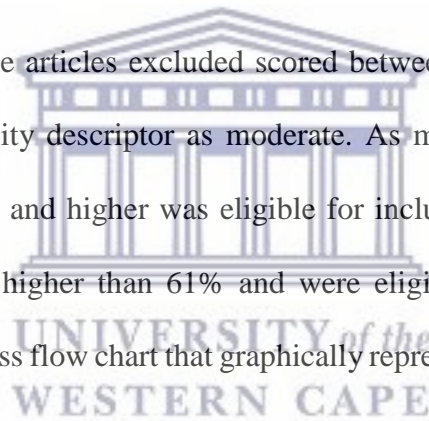
The 112 titles that were deemed relevant during the previous step were screened using the stipulated inclusion and exclusion criteria. After the screening, 66 abstracts were excluded. The primary reasons for the exclusion of articles were that they did not include the assessment tools for the assessment of adult ADHD. Other reasons for exclusion included children as participants, various treatment strategies of ADHD, and studies related to ADHD and its association to other phenomena (e.g. other disorders).

Ten abstracts were identified as requiring additional information. These abstracts provided insufficient information to determine whether it qualified for inclusion. Upon closer inspection seven records were excluded. Five of the seven excluded abstracts contained

assessment tools assessing aspects of adult ADHD, such as quality of life and levels of functional impairment. The instruments in these articles were not screening or diagnosing for adult ADHD as stipulated in the inclusion criteria. The other two articles were excluded as it included children as participants. The remaining three articles were included with the articles selected for full text appraisal. Thus, a combined total of 39 records were included and progressed to the next step.

### **Step 3: Eligibility**

Of the 39 articles, eight articles were not available in full text on the UWC databases and were excluded on this basis. The remaining 31 articles were critically appraised. As mentioned before the threshold score was 60%. Five articles were excluded as it did not exceed the threshold requirement. The articles excluded scored between 41 and 60% in the full text appraisal and obtained a quality descriptor as moderate. As mentioned before, articles that obtained a total score of 61% and higher was eligible for inclusion in the final review. The remaining 26 articles scored higher than 61% and were eligible for inclusion in the final summation. Figure 2 is a process flow chart that graphically represents the information reported above.



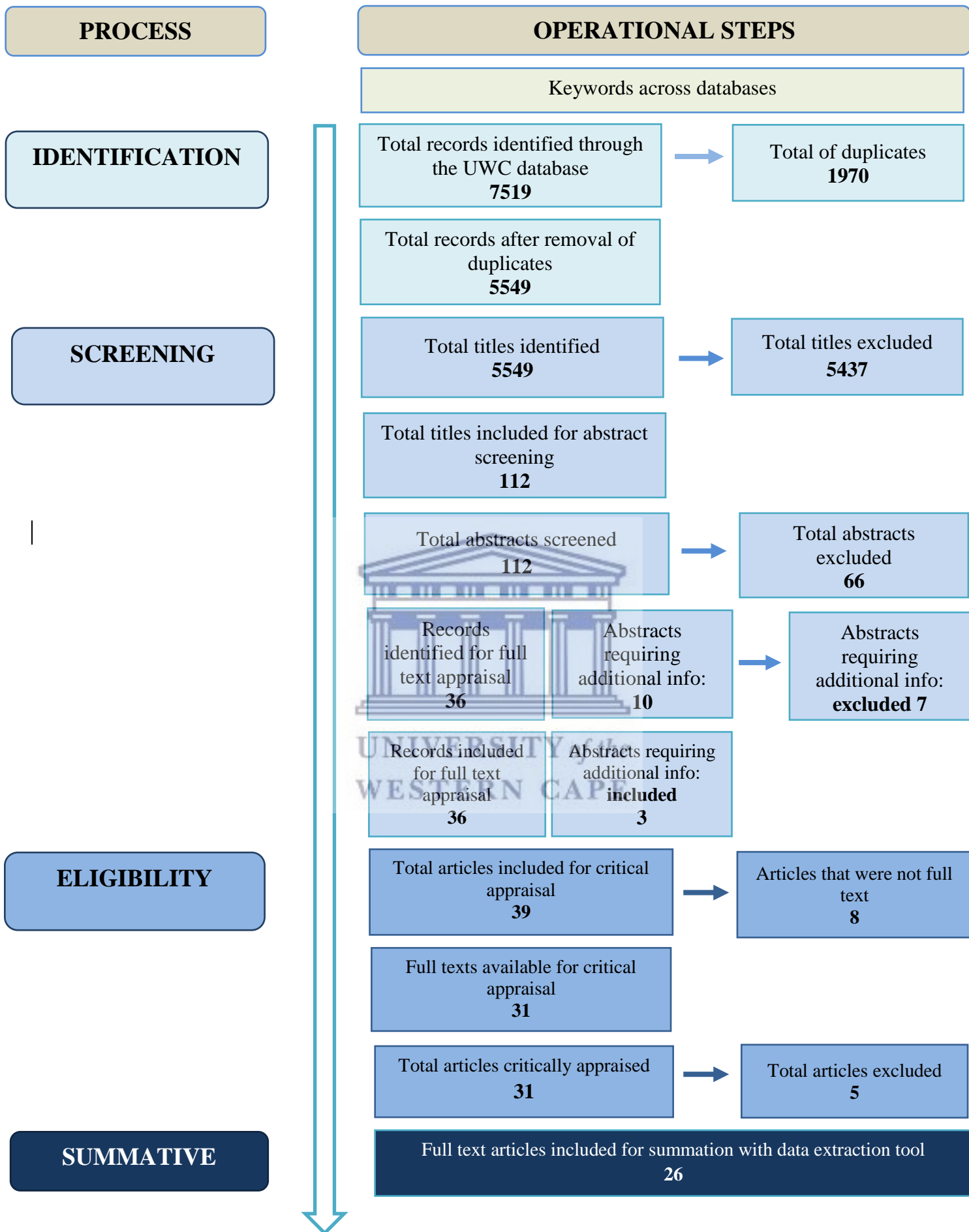


Figure 4.1: Review process

## 4.2. Descriptive Meta-synthesis:

The Descriptive Meta-Synthesis included two sections, namely, ranking, and a summary of data extracted from articles included in the final review.

### 4.2.1 Ranking

The articles were ranked in descending order based on their overall score obtained on the critical appraisal tool. Thus, the higher ranks reflect articles that scored higher on methodological rigour as assessed by the SFS scoring system. Table 4.1 below summarizes the ranked scores per article.

**Table 4.1: Ranking**

Rank no.	Authors	Appraisal Category	Final Appraisal Score
1	Takeda et al 2015	Excellent	84%
2	Amandor-Campos et al 2016	Excellent	82%
2	Kim et al 2013	Excellent	82%
2	Evren et al 2016	Excellent	82%
3	Kingston et al 2013	Excellent	81%
4	Dvorsky et al 2016	Strong	79%
4	Gorlin et al 2016	Strong	79%
5	Ramos-Quiroga et al 2015	Strong	78%
5	Manor et al 2012	Strong	78%
6	Singh et al 2015	Strong	75%
6	Vidal et al 2014	Strong	75%
7	Christiansen et al 2012	Strong	74%
7	Eich et al 2012	Strong	74%
7	Fuller-Killgore et al 2012	Strong	74%
7	Mórtbert et al 2012	Strong	74%
8	Faries et al 2012	Strong	73%
9	Gray et al 2014	Strong	72%
9	Ramos-Quiroga et al 2012	Strong	72%
9	Amandor-Campos et al 2014	Strong	72%
9	Daigre et al 2015	Strong	72%
9	Van der Glind et al 2013	Strong	72%
9	Young et al 2016	Strong	72%
9	Marshall et al 2016	Strong	72%
10	Marchant et al 2013	Strong	68%
10	Söderström et al 2014	Strong	68%
11	Kooij et al 2013	Strong	66%

The majority (21) of the articles eligible scored in the strong category (61-80%). The remaining five scored in the excellent category (81-100%). The study conducted by Takeda et al (2015) was ranked first. This study obtained the highest appraisal score of 84%. The following is a breakdown of how articles scored in the respective sections and subsections of the appraisal tool.

#### **4.2.2 Ranks based on subsections**

The critical appraisal was comprised of two sections as mentioned before. The ranking of the articles will be discussed per subsection.

##### **4.2.2.1 Section A**

Overall, articles generally scored well in section A of the appraisal tool. The maximum score that an article could achieve in this section was a total score of 44. This 44 was comprised of smaller scores that studies attained in the various subsections. The Table 4.2 below summarises how each study ranked in the subsections. It also includes the final appraisal score, which indicates how each article ranked in the full appraisal. This was included to determine whether the top ranked articles generally scored well in section A. Articles in the excellent category for the final appraisal score generally scored well for section A of the appraisal. A discussion of how these articles ranked for section A and across subsections is discussed below.

**Table 4.2: Section A**

Authors	Overall Ranking	Section A Ranking	Purpose (Max score = 7)	Design (Max score = 4)	Sampling Type (Max score = 9)	Data Collection (Max score = 6)	Data Analysis (Max score = 4)	Results (Max score = 12)	Ethics (Max score = 2)
Takeda et al 2015	1	2 (42)	1 (7)	2 (3)	2 (8)	1 (6)	1 (4)	1 (12)	1 (2)
Amandor-Campos et al 2016	2	3 (40)	1 (7)	3 (2)	1 (9)	1 (6)	1 (4)	2 (11)	2 (1)
Kim et al 2013	2	4 (39)	2 (6)	3 (2)	3 (7)	1 (6)	1 (4)	1 (12)	1 (2)
Evren et al 2016	2	1 (44)	1 (7)	1 (4)	1 (9)	1 (6)	1 (4)	1 (12)	1 (2)
Kingston et al 2013	3	3 (40)	2 (6)	1 (4)	2 (8)	1 (6)	1 (4)	3 (10)	1 (2)
Dvorsky et al 2016	4	5 (38)	1 (7)	2 (3)	4 (6)	1 (6)	1 (4)	3 (10)	1 (2)
Gorlin et al 2016	4	1 (44)	1 (7)	1 (4)	1 (9)	1 (6)	1 (4)	1 (12)	1 (2)
Ramos-Quiroga et al 2015	5	2 (42)	1 (7)	1 (4)	1 (9)	1 (6)	1 (4)	3 (10)	1 (2)
Manor et al 2012	5	2 (42)	2 (6)	1 (4)	1 (9)	1 (6)	1 (4)	2 (11)	1 (2)
Singh et al 2015	6	5 (38)	1 (7)	1 (4)	3 (7)	1 (6)	1 (4)	5 (8)	1 (2)
Vidal et al 2014	6	5 (38)	1 (7)	3 (2)	3 (7)	1 (6)	1 (4)	3 (10)	1 (2)
Christiansen et al 2012	7	3 (40)	3 (5)	1 (4)	3 (7)	1 (6)	1 (4)	1 (12)	1 (2)
Eich et al 2012	7	2 (42)	2 (6)	1 (4)	2 (8)	1 (6)	1 (4)	1 (12)	1 (2)
Fuller-Killgore et al 2012	7	6 (37)	1 (7)	3 (2)	3 (7)	1 (6)	1 (4)	3 (10)	2 (1)
Mórtbert et al 2012	7	7 (36)	2 (6)	3 (2)	2 (8)	1 (6)	1 (4)	4 (9)	2 (1)
Faries et al 2012	8	8 (35)	3 (5)	1 (4)	2 (8)	1 (6)	1 (4)	5 (8)	1 (2)
Gray et al 2014	9	4 (39)	1 (7)	1 (4)	3 (7)	1 (6)	1 (4)	4 (9)	1 (2)
Ramos-Quiroga et al 2012	9	3 (40)	1 (7)	1 (4)	3 (7)	1 (6)	1 (4)	3 (10)	1 (2)
Amandor-Campos et al 2014	9	7 (36)	1 (7)	3 (2)	1 (9)	2 (3)	1 (4)	3 (10)	2 (1)
Daigre et al 2015	9	5 (38)	1 (7)	1 (4)	3 (7)	1 (6)	1 (4)	5 (8)	1 (2)
Van der Glind et al 2013	9	6 (37)	2 (6)	1 (4)	3 (7)	1 (6)	1 (4)	5 (8)	1 (2)
Young et al 2016	9	4 (39)	2 (6)	1 (4)	3 (7)	1 (6)	1 (4)	3 (10)	1 (2)
Marshall et al 2016	9	5 (38)	1 (7)	1 (4)	4 (6)	1 (6)	1 (4)	3 (10)	2 (1)
Marchant et al 2013	10	9 (34)	4 (4)	3 (2)	3 (7)	2 (3)	1 (4)	1 (12)	1 (2)
Söderström et al 2014	10	6 (37)	2 (6)	2 (3)	2 (8)	1 (6)	1 (4)	4 (9)	2 (1)
Kooij et al 2013	11	8 (35)	2 (6)	3 (2)	5 (5)	1 (6)	1 (4)	3 (10)	1 (2)

### ***a) Purpose***

The purpose subsection of the appraisal tool determined whether studies provided adequate information on the context and research problem. The 26 articles were distributed across four ranks. This truncated ranking suggested that the studies scored very similarly in this section. Fourteen studies scored 44 points and were ranked joint first (1<sup>st</sup>). Nine articles scored 42 and were ranked second (2<sup>nd</sup>). Two articles scored 40 and were ranked third (3<sup>rd</sup>). The articles ranked within the first three positions scored highly in this subsection which meant that they provided information consistent with that which is expected of good reporting. Readers were orientated with a comprehensive background and the nature of the problem was clearly articulated in line with the requirements of journals for publication.

The remaining rank all scored between 30 and 40 in this subsection. These articles were less consistent in their formulation of the research problem and provided brief contextualization of the study. The scores in this subsection suggest that there is a clear expectation of what should be reported in manuscripts. Authors then prioritize adherence to format and the instructions of journals, as this has been cited as main reasons why manuscripts are rejected (Whitehouse, 2013). Hence these truncated scores potentially reflect publication bias.

### ***b) Design***

Studies were appraised to assess how the research design undertaken in respective studies was reported. Appraisers also sought to determine if the design incorporated was appropriate and applicable to address the aims of the particular studies. The scores of the articles were narrowly distributed across three ranks. Fifteen studies were ranked joint first (1<sup>st</sup>). They occupied this rank as they achieved the highest score possible (7) in this subsection. Three studies scored six and were ranked second (2<sup>nd</sup>), and eight studies scored five and were ranked third (3<sup>rd</sup>). The narrow band of ranks and relatively high scores were indicative of the authors making intentional decisions to prioritize thorough descriptions of the research designs of their

respective studies. This manner of detailed reporting contributes to the replicability of the studies. Replicability is an important characteristic of good quality reporting and by extension good quality research.

### ***c) Sample***

The subsection on sampling evaluated whether the authors of respective studies reported on the sampling methods and procedures undertaken. This section also sought to evaluate whether the methods were suitable given the research aims. The articles were distributed across five ranks. Six articles obtained full marks (9) thereby were ranked joint first (1<sup>st</sup>). Six studies scored eight for this subsection therefore ranked second (2<sup>nd</sup>). Eleven studies ranked third with a score of seven (3<sup>rd</sup>). These articles all reported in detail on the sampling strategies. Two articles were ranked fourth (4<sup>th</sup>) and scored six points. These articles generally scored less, because they failed to identify the type of sampling strategy and whether it constituted probability or non-probability sampling. Only one study did not report on sampling recruitment and subsequently ranked fifth (5<sup>th</sup>) with a lower score in this subsection. Academic and publication conventions include reporting on sampling strategies and information on the samples included within studies. The top four ranks reported on sampling in a manner that was consistent with these conventions.

### ***d) Data Collection***

The data collection subsection evaluated how articles reported on the methods used to collect data and the suitability thereof. Studies generally performed well and the ranking was truncated. Articles were distributed across only two ranks. A total of 24 articles obtained perfect scores (6) and were ranked joint first (1<sup>st</sup>). These articles scored well and included a clear motivation for the selection of the data collection methods, as well as a good description of instrumentation i.e. how the instruments were used. This made it more readily apparent that the methods of data collection were appropriate for the stated aims of the study and higher



scores could be allocated. The remaining two studies scored 3 and were ranked joint second (2<sup>nd</sup>). These articles generally were less clear on the motivation for using particular methods for data collection and therefore scored lower marks.

Calfee and Valencia (2010) emphasised that it was necessary to include thorough descriptions of the methods and procedures utilized within the method section. This is a necessary requirement for publication as well. These truncated scores reflect two distinct groups. One group reported in detail on their methodological decisions including motivations and the other merely indicated which methods were used. Studies that scored well in this section provided sufficient detail that would assist in replication.

#### *e) Data Analysis*

The data analysis subsection appraised whether the data analysis was described in detail and secondly, whether the data analysis was appropriate relative to the research question. In this subsection, all studies achieved perfect scores and were ranked joint first (1<sup>st</sup>). It is apparent that the data analysis methods of these respective studies were clearly identified and motivated. Intervention studies using quantitative methods tend to have more detailed reporting on sampling, design and analysis sections. This is typical of quantitative methods as the results are contingent on these processes. There is a clear expectation of the kind of detail that is required when reporting quantitative findings. This lack of variation in the scoring was indicative of adherence to the academic and publication conventions. Thus, it reflects those biases. It indicates that authors emphasize certain sections over others to follow specific conventions, thereby increasing the likelihood of being published.

### ***f) Results and Conclusion***

This subsection was concerned with the evaluation of a number of factors related to the reporting of the results and findings. It included an assessment on how clearly and unambiguously the findings, conclusions and limitations were reported. Studies were distributed across five ranks. Seven studies achieved perfect scores and were ranked joint first (1<sup>st</sup>). Two scored eleven and ranked second (2<sup>nd</sup>). The first and second ranked articles scored well and presented information on findings, limitations and recommendations clearly. The data and results clearly supported the discussion, conclusion and recommendations. Ten studies were ranked third (3<sup>rd</sup>), and three were ranked fourth (4<sup>th</sup>). These articles scored lower on this subsection. The reporting here was not clearly supported by the findings, or were unclear and poorly constructed. For example, where statistical significance was present, the alpha levels were not reported. One study obtained the lowest mark and was ranked fifth (5<sup>th</sup>). This article omitted reporting on statistical significance.

### ***g) Ethics***

This section of the appraisal tool assessed whether the authors reported that they received ethics approval for conducting their research and whether they identified the issuing ethics committee. The articles were distributed across two ranks only. Twenty studies reported fully on this and were ranked joint first (1<sup>st</sup>). These studies scored optimally in this section, however, a tendency not to report on the specific ethical principles was noted. Six articles were ranked second (2<sup>nd</sup>). They scored lower in this section, as only one sentence stating that the study was approved by the relevant ethical institutions was included. This might be a function of publication as many journals require that ethics certificates are uploaded when the article is submitted. Thus, the authors assume that this is sufficient and opt not to report in detail on how ethics principles were applied to the respective studies. The decision to report minimally in this section is also influenced by the stringent word count requirements of journals that might

influence what authors report. The clear trend observed across all 26 articles was that ethics as a subsection was reported on more generally as opposed to other subsections.

#### **4.2.2.2 Section B**

Section B of the appraisal tool was concerned with instrumentation. The maximum score that an article could achieve in section B was a total score of 51. This score was a composite score that included various subsections. Table 4.3 below summarises how each study ranked in the respective subsections. It also includes the final appraisal score, which indicates how each article ranked in the full appraisal. This was included to compare how the top ranked articles overall compared to the ranking in section B. Articles in the excellent category for the final appraisal score generally scored well for section B of the appraisal. A discussion of how these articles ranked for section B and across subsections is discussed below.



**Table 4.3 Section B**

Authors	Overall Ranking	Purpose (Max score = 10)	Methodology (Max core = 6)	Items & Administration (Max score = 18)	Reliability & Validity (Max score = 10)	Interpretation (Max score = 7)
Takeda et al 2015	1	1 (10)	1 (6)	2 (12)	3 (7)	2 (4)
Amandor-Campos et al 2016	2	1 (10)	1 (6)	1 (13)	4 (6)	3 (3)
Kim et al 2013	2	1 (10)	1 (6)	3 (11)	2 (8)	3 (3)
Evren et al 2016	2	2 (9)	2 (5)	5 (9)	2 (8)	4 (2)
Kingston et al 2013	3	2 (9)	1 (6)	3 (11)	2 (8)	3 (3)
Dvorsky et al 2016	4	1 (10)	2 (5)	4 (10)	1 (9)	4 (2)
Gorlin et al 2016	4	1 (10)	3 (4)	6 (8)	1 (9)	3 (3)
Ramos-Quiroga et al 2015	5	2 (9)	3 (4)	6 (8)	2 (8)	4 (2)
Manor et al 2012	5	1 (10)	1 (6)	4 (10)	6 (4)	4 (2)
Singh et al 2015	6	2 (9)	3 (4)	1 (13)	6 (4)	3 (3)
Vidal et al 2014	6	1 (10)	1 (6)	4 (10)	3 (7)	4 (2)
Christiansen et al 2012	7	4 (7)	4 (3)	5 (9)	2 (8)	3 (3)
Eich et al 2012	7	2 (9)	4 (3)	6 (8)	5 (5)	3 (3)
Fuller-Killgore et al 2012	7	1 (10)	1 (6)	6 (8)	2 (8)	4 (2)
Mórtbert et al 2012	7	1 (10)	1 (6)	6 (8)	3 (7)	4 (2)
Faries et al 2012	8	1 (10)	4 (3)	4 (10)	6 (4)	1 (5)
Gray et al 2014	9	3 (8)	1 (6)	6 (8)	6 (4)	4 (2)
Ramos-Quiroga et al 2012	9	3 (8)	3 (4)	6 (8)	4 (6)	4 (2)
Amandor-Campos et al 2014	9	1 (10)	4 (3)	4 (10)	3 (7)	4 (2)
Daigre et al 2015	9	2 (9)	4 (3)	4 (10)	5 (5)	3 (3)
Van der Glind et al 2013	9	1 (10)	3 (4)	6 (8)	6 (4)	1 (5)
Young et al 2016	9	1 (10)	2 (5)	6 (8)	6 (4)	3 (3)
Marshall et al 2016	9	1 (10)	2 (5)	6 (8)	6 (4)	3 (3)
Marchant et al 2013	10	2 (9)	3 (4)	6 (8)	3 (7)	3 (3)
Söderström et al 2014	10	2 (9)	4 (3)	6 (8)	6 (4)	3 (3)
Kooij et al 2013	11	3 (8)	3 (4)	5 (9)	5 (5)	4 (2)

### ***a) Purpose of Instrument***

This section of the appraisal tool was concerned with a number of aspects associated with reporting on the purpose of the instrument in question. This included whether authors discussed key characteristics of the tools and identified it as screening instruments or in-depth diagnostic instruments. Scores were distributed across four ranks. Fourteen articles obtained maximum scores (10) and were ranked joint first (1<sup>st</sup>) in this subsection. Eight studies scored nine and were ranked joint second (2<sup>nd</sup>). Two studies scored eight and were ranked joint third (3<sup>rd</sup>). The top three ranks were similar, and these articles generally included sections that briefly introduced the instruments used in the respective studies, and hence the truncated ranking. The detailed reporting on the instrument was important since these instruments were used as the primary means of collecting data.

One article obtained a slightly lower score (6) and was ranked fourth (4<sup>th</sup>). This article did not specify what characteristics could be measured and what decisions could be made based on the particular assessments. The instrument used in this study was well established and widely used. Thus, it appears that the authors opted not to report on the purpose of the instrument as they accepted that there was consensus in the fraternity on the purpose and standing of this instrument. This reflects how within conventions of clinical work and research, there are forms of consensual agreement that serve as validation for certain methodological or clinical decisions.

### ***b) Methodology***

The methodology subsection determined whether studies defined the construct of ADHD in adults, theoretically and operationally. The scores were distributed across four ranks. Nine articles obtained optimal scores (6) for this subsection and were ranked joint first (1<sup>st</sup>). Four articles obtained five (5) and were ranked second (2<sup>nd</sup>), while seven studies obtained four (4) and were ranked third (3<sup>rd</sup>). These studies scored similarly and generally reported on the

definition of the construct adult ADHD. Six articles obtained three (3) and were ranked fourth (4th). These studies scored lower and only obtained half of the possible score in this subsection. These articles did not clearly report on the definition of the construct adult ADHD, domains included, nor how these would be measured.

### ***c) Items and Administration***

This subsection evaluated whether authors discussed the administration of the instruments. Additionally, the reporting on the construction and standardization of the administration procedures were assessed. The studies were distributed across six ranks. Two articles obtained the highest score of 13 and were ranked joint first (1<sup>st</sup>). Only one study ranked second (2<sup>nd</sup>) with a score of 12 out of a possible 18. Two studies recorded a score of 11 and were ranked third (3<sup>rd</sup>). Six studies ranked fourth (4<sup>th</sup>) with a score of 10. Two studies obtained 9 and were ranked fifth (5<sup>th</sup>). Twelve studies scored 8, therewith ranking sixth (6<sup>th</sup>). Studies scoring twelve and less in this subsection consistently did not report on the administration procedures necessary for the successful completion of the instruments. This pattern or consistent nature of not reporting proposes that it may not have been necessary for publication. These studies were intervention studies where the instrument was used to indicate symptomology as an output and inclusion criterion. Thus, the focus was not on the instrumentation, but the clinical intervention. The highest ranked article specifically reported on the psychometric properties of the instrument, and was expected to report on matters related to construction. Thus, it appears that the information reported on for this subsection was influenced by publication bias or academic convention.

### ***d) Reliability and Validity***

The reliability and validity subsection of the appraisal tool was included to evaluate the articles based on whether they presented information on the psychometric properties of the instruments. Scores were distributed across six ranks. Two studies scored nine and were ranked

first (1<sup>st</sup>). Five studies obtained eight out of a possible 10 and were ranked second (2<sup>nd</sup>). Five articles scored seven and were ranked third (3<sup>rd</sup>). Two scored six and were ranked fourth (4<sup>th</sup>).

Three scored five and were ranked fifth (5<sup>th</sup>). Twelve studies scored four and were ranked sixth (6<sup>th</sup>). Studies that scored less than half of the possible scores in this subsection did not mention any forms of validity. Often studies reported on the reliability and validity measured in previous studies, thus they did not calculate these properties themselves. A possible reason for this is that the objective was often to measure adult ADHD using instruments that report good psychometric properties. This differed to studies that tested the reliability and validity of the respective instruments, as the objective here was to determine the psychometric properties for measuring adult ADHD in specific contexts. Finally, articles often reported on the sensitivity and specificity or the predictive values of the instruments rather than reliability and validity. In this way, authors were concerned with reporting on the accuracy of the instruments in successfully measuring adult ADHD or successfully not measuring adults without ADHD. The focus was thus on higher level attributes rather than reliability and validity that has been reasonably well established in the literature. This manner of reporting thus reflects current trends in instrumentation and psychometric construction and ultimately reflects academic convention (Foxcroft, 2011; Laher & Cockcroft, 2013).

#### ***e) Interpretation***

This subsection assessed the extent to which authors demonstrated an awareness of other instruments measuring similar constructs. The articles were also evaluated in terms of whether there was any reference to the guide for interpreting the scores of the instruments. The majority of the articles scored lower in this subsection. Studies were distributed across four ranks. Articles generally scored lower as authors generally provided a description of the tools but did not report on the development of the instrument. It is clear that such information was not prioritized. Authors focused on the scores generated in the context of their studies and whether

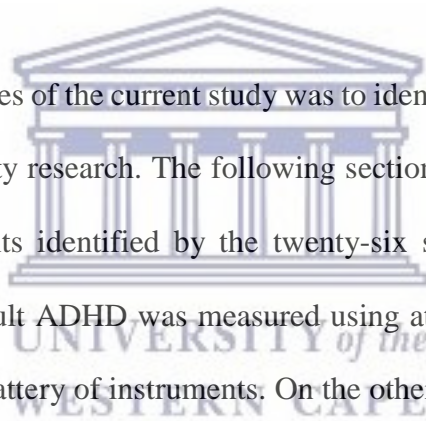
significant effects were observed. The omission of more detailed information on the interpretation was accepted or known practice.

#### **4.2.3 Summary**

A total of twenty-six articles were included in the final summation as they met the threshold score of 60%. Articles generally scored higher on section A than on section B. One of the reasons for this is that authors prioritized reporting on information necessary for publication, and in doing so, reported less on instrumentation. This might be indicative of academic convention and publication bias that in turn detracts from the methodological coherence and rigour of the article and adversely impacts replication.

#### **4.3 Instruments**

One of the main objectives of the current study was to identify instruments used to assess adult ADHD from good quality research. The following section contains a brief summary of the sixteen distinct instruments identified by the twenty-six studies included for the final summation. In each study, adult ADHD was measured using at least one of the instruments, and some studies included a battery of instruments. On the other hand, certain studies utilized the same measure, albeit in different contexts. Table 4.4 below illustrates the instruments used in the respective studies.





**Table 4.4 Instruments measuring ADHD**

<b>Authors</b>	<b>Instruments</b>
Amandor-Campos et al 2014	CAARS
Amandor-Campos et al 2016	CAARS-L
Christiansen et al 2012	CAARS
Daigre et al 2015	ASRS Screener
Dvorsky et al 2016	BAARS-IV
Eich et al 2012	SCL-90-R
Evren et al 2016	ASRS-v1.1
Faries et al 2012	PDI-4
Fuller-Killgore et al 2012	ASRS CAARS Brown ADD
Gorlin et al 2016	SCID-5
Gray et al 2014	ASRS
Kim et al 2013	ASRS
Kingston et al 2013	ASRS CAARS WURS Brown ADD scale IVA+Plus
Kooij et al 2013	CAARS (interrogated V)
Manor et al 2012	ADHD-SQ
Marchant et al 2013	WRAADDs
Marshall et al 2016	BAARS-IV
Mórtbert et al 2012	WURS
Ramos-Quiroga et al 2012	CAADID
Ramos-Quiroga et al 2015	PRISM
Singh et al 2015	IPDE-SQ
Söderström et al 2014	CSS
Takeda et al 2015	ASIA
Van der Glind et al 2013	ASRS
Vidal et al 2014	CAADID
Young et al 2016	BAARS-IV

The following section provides information on the identified instruments. Three core issues were reported on, namely, the arrangement of the instrument including the type, format, and purpose of the respective measures.

### 4.3.1 Types of instruments

The instruments identified differed in the type of disorder they measured. Table 4.5 lists the different types of instruments used to identify adult ADHD by the respective studies.

**Table 4.5 Type of instrument**

Name of Instrument	Types or focus
Attention Deficit Hyperactivity Disorder Screening Questionnaire (ADHD-SQ)	Adult ADHD
Assessment System for Individuals with ADHD (ASIA)	Adult ADHD
Adult ADHD Self-Report Scale (ASRS)	Adult ADHD
Barkley Adult ADHD Self-Report Forms-IV(BAARS-IV)	Adult ADHD
Brown ADD scale (Brown ADD scale)	Adult ADHD
Conners Adult ADHD Diagnostic Interview for DSM IV-TR (CAADID)	Adult ADHD
Conners Adult ADHD Rating Self-Report (CAARS)	Adult ADHD
Current Symptom Scale (CSS)	Adult ADHD
International Personality Disorder Examination (IPDE-SQ)	Personality Disorders
Integrated Visual and Auditory Continuous Performance Test (IVA+Plus)	ADHD
The Provisional Diagnostic Instrument (PDI-4)	MDE; GAD; ADHD; Bipolar 1
Psychiatric Research Interview for Substance and Mental disorders-IV (PRISM)	Substance & Mental Disorders
Structural Clinical Interview for the DSM-5 (SCID-5)	General DSM-5
Symptom Checklist 90 Revised (SCL-90-R)	Multidimensional
Wender-Reimherr Adult Attention Deficit Disorder Scale (WRAADDS)	Adult ADHD
Wender-Utah Rating Scale (WURS)	Adult ADHD

The identified instruments consisted of two categories; 1) instruments specifically developed for measuring adult ADHD; and 2) diagnostic instruments developed for measuring personality and other multidimensional disorders.

- 1) ADHD specific instruments: Eleven instruments were specifically designed for measuring ADHD. Ten of which, namely, the ADHD-SQ; ASIA; ASRS; BAARS-IV; Brown ADD scale; CAADID; CAARS; CSS; WRAADDS; and WURS are adult ADHD measures. The IVA+Plus is, on the other hand, a screening tool used for measuring ADHD in both child and adult populations.

2) Diagnostic instruments: Five measures were diagnostic instruments not specific to adult ADHD, namely, the IPDE-SQ; PRISM; PDI-4; SCID-5; SCL-90-R. These instruments were standardized for measuring or assessing other psychological constructs. For instance, the IPDE-SQ is considered as the gold standard tool for the diagnosis of personality disorders. The PRISM was specifically developed to measure comorbidity in substance use disorders. The PDI-4 is an instrument utilized in the determination of four psychiatric disorders in primary health care settings. These four disorders are MDE; GAD; ADHD; and Bipolar 1. Similarly, the SCID-5 and SCL-90-R are multidimensional instruments associated and used to assess a number of different diagnostic syndromes including, but not limited to, ADHD. The SCID-5 was adapted to reflect the diagnostic changes in the DSM-5 diagnostic criteria. These instruments were included in studies to determine the applicability of diagnosing ADHD in adult psychiatric clients. Studies therefore reported on the psychometric properties of these instruments.

#### **4.3.2 Format of Instruments**

Table 4.6 reflects the different formats of the identified instruments.

**Table 4.6 Format of identified instruments**

<b>Name of Instrument</b>	<b>Format</b>
Attention Deficit Hyperactivity Disorder Screening Questionnaire (ADHD-SQ)	Self-Report
Assessment System for Individuals with ADHD (ASIA)	Semi-Structured Interview
Adult ADHD Self-Report Scale (ASRS)	Self-Report
Barkley Adult ADHD Self-Report Forms-IV(BAARS-IV)	Self + Other Report
Brown ADD scale (Brown ADD scale)	Self-Report
Conners Adult ADHD Diagnostic Interview for DSM IV-TR (CAADID)	Semi-Structured Interview
Conners Adult ADHD Rating Self-Report (CAARS)	Self + Observer-Report
Current Symptom Scale (CSS)	Self-Report
International Personality Disorder Examination (IPDE-SQ)	Self-Report
Integrated Visual and Auditory Continuous Performance Test (IVA+Plus)	Computerized Continuous Performance Test
The Provisional Diagnostic Instrument (PDI-4)	Self-Report
Psychiatric Research Interview for Substance and Mental disorders-IV (PRISM)	Structured interview
Structural Clinical Interview for the DSM-5 (SCID-5)	Structured Interview
Symptom Checklist 90 Revised (SCL-90-R)	Self-Report
Wender-Reimherr Adult Attention Deficit Disorder Scale (WRAADDS)	Clinician Rated scale
Wender-Utah Rating Scale (WURS)	Retrospective Self-Report

Illustrated in Table 4.6, instruments used for measuring the construct of adult ADHD either took on the form of self-report, observer-report, or clinician-administered scales (Epstein & Kollins, 2006; Marchant et al, 2013). Ten of the instruments identified in this review were self-report instruments. These included, the ADHD-SQ; ASRS; BAARS-IV; Brown ADD scale; CAARS; CSS; IPDE-SQ; PDI-4; SCL-90-R; and WURS. Authors of these instruments recognised that self-report measures definitely provide valuable information in the identification of adult ADHD symptoms. Similarly, the WURS is a retrospective self-report measure, which is completed by adults to assess their childhood history of ADHD symptoms.

Two studies, Kingston et al (2013) and Mrtberg et al (2012) included the WURS, which, in both studies, was administered with other self-report instruments. The criticisms against self-

reports and retrospective instruments are offset by combining such instruments with companion measures. The findings suggest that in good quality research instruments are used in tandem with other known measures rather than independently. In this way, the findings are triangulated with the use of multiple measures that in turn enhances the methodological rigour and coherence of the study.

Only one of the instruments identified was an observer-report instrument. This is the CAARS instrument which includes both self-report and observer-report forms. The observer-report instrument enables family and friends to provide collateral information regarding the identification of adult ADHD symptoms in the patient's lives. Studies that included the observer-report for adult ADHD often used the observer-report as an additional instrument, in conjunction with self-report measures. In this way, observer-reports are valuable as observers, and patients are able to identify symptoms for adult ADHD with collateral. This has a similar effect to the use of companion measures reported above.

Five of the instruments listed were formatted as clinical interviews. Clinician rated scales take on the form of interviews. Interviews enable clinicians to attain relevant information from the patient or informant for diagnosis (Marchant et al, 2013). The SCID-5 and the PRISM were structured interviews. The AISA, the CAADID and WRAADS were constructed as semi-structured interviews.

### **4.3.3 Versions of instruments**

Another distinction in form that became evident was that instruments had various versions of the same instrument. Table 4.7 lists the instruments that specified various versions.

**Table 4.7 Versions of instruments**

Name of Instrument	Versions	Items
ASRS	Short Version	6 items
	Long Version	18 items
CAARS (Self/Observer report)	Short Version	26 items
	Long Version	66 items
	Investigator report Version	30 items
IPDE-SQ	Different versions not specified	-
PRISM	Different versions not specified	-
SCID-5	Different versions not specified	-

A number of the instruments identified in this study have different versions of the same instrument for measuring ADHD. The intention here was not to discuss all the different versions of the specific instruments, rather to encapsulate only the versions identified by the articles included in the final summation. In the table above five instruments specified multiple versions, namely, the ASRS; CAARS; IPDE-SQ; PRISM; SCID-5.

The ASRS-v1.1 has two different versions. These versions are referred to as the short version and long version. The main difference between these versions are in terms of the number of items. The short version has six items and the long version is comprised with the same six items as well as an additional 12 items accounting to a total of eighteen items. Importantly, the authors recognised the shortened version of the ASRS-v1.1 as most predictive of symptoms associated with ADHD. The 6 items of the short version are considered to be sufficient to screen for adult ADHD. The Part B of the long version only provides additional clinical information to determine the level of impairment.

Another instrument that differentiates between the short version and long version is the CAARS. The short form is comprised with 26 items and the longer version has 66 items. Although the shorter forms have less items it assesses similar symptoms however to a lesser

extent. In other words, where the long version has 12 items within a particular subscale, the shorter version has 5 items for that same subscale. Shorter versions are therefore prioritized as a quick screener of adult ADHD taking approximately 10 minutes.

Moreover, this instrument further distinguishes between self-report version, observer report version, and the investigator version. Both the self-report and the observer-report has a long version and a short version, and the distinction of these scales is the manner in which the questions are phrased. Although the items are essentially worded and ordered the same, the self-report is formulated in the first person and the observer-reports in the third person (Amador-Campos et al, 2012). The investigator version of the CAARS is different in word and order. It has 30 items formulated in a manner that allows the clinician to screen for symptoms associated with adult ADHD.

The other instruments that reported other versions were instruments that were not specific to adult ADHD, rather they were instruments well known for measuring personality disorders (IPDE-SQ), DSM-5 disorders (SCID-5), or comorbid with substance disorders (PRISM). It is apparent that well known instruments have different versions of the same instrument. An obvious rational for having a different version in length is the time it takes to administer the screening. Shorter versions require less time but also provide less information. In this manner an advantage of longer measures is that it provides more information on the clients and their symptoms. A benefit of having a self-report, observer-report and investigator report is that a client is afforded an option of completing a battery of instruments. In fact, it is largely considered best practice to make use of a combination of self-report, observer-report and clinician administered instruments (Kingston et al, 2013). Hence these scales become more attractive for having various versions.

Similarly, well known instruments have often been adapted across a number of languages. Researchers have often sought to translate these instruments into other languages as opposed to formulating new instruments relevant to specific languages and cultures. Arguably, instruments that report more than one version in length, language and type of report become more convenient and accessible than those who do not.

#### 4.3.4 Purpose of Instrument

There are different types of instruments associated with measuring adult ADHD. These instruments are constructed and structured differently, often according to the specific purpose.

Table 4.8 lists the purposes of the instruments identified.

**Table 4.8 Purpose of the instrument**

Name of Instrument	Purpose
Attention Deficit Hyperactivity Disorder Screening Questionnaire (ADHD-SQ)	Screening Instrument
Assessment System for Individuals with ADHD (ASIA)	Diagnostic Instrument
Adult ADHD Self-Report Scale (ASRS)	Screening Instrument
Barkley Adult ADHD Self-Report Forms-IV(BAARS-IV)	Screening Instrument
Brown ADD scale (Brown ADD scale)	Screening Instrument
Conners Adult ADHD Diagnostic Interview for DSM IV-TR (CAADID)	Diagnostic Instrument
Conners Adult ADHD Rating Self-Report (CAARS)	Screening Instrument
Current Symptom Scale (CSS)	Screening Instrument
International Personality Disorder Examination (IPDE-SQ)	Screening instrument
Integrated Visual and Auditory Continuous Performance Test (IVA+Plus)	Screening Instrument
The Provisional Diagnostic Instrument (PDI-4)	Diagnostic Instrument
Psychiatric Research Interview for Substance and Mental disorders-IV (PRISM)	Diagnostic Instrument
Structural Clinical Interview for the DSM-5 (SCID-5)	DSM-5 Diagnosis
Symptom Checklist 90 Revised (SCL-90-R)	Screening instrument
Wender-Reimherr Adult Attention Deficit Disorder Scale (WRAADDS)	Diagnostic Instrument
Wender-Utah Rating Scale (WURS)	Screening instrument

In Table 4.8 above, ten instruments listed (ADHD-SQ; ASRS; BAARS-IV; Brown ADD Scale; CAARS; CSS; IPDE-SQ; IVA+Plus; SCL-90-R; and WURS) were identified as screening instruments. These are instruments that serve the purpose of testing only for the



presence of adult ADHD symptoms. Conversely, the remaining six instruments (AISA; CAADID; PDI-4; PRISM; SCID-5; and WRAADDs) were referred to as diagnostic instruments. These are instruments used for the purpose of assessing for the diagnosis of adult ADHD.

Interestingly, the instruments identified as screening for the presence of adult ADHD symptoms were all constructed as self- or observer-report forms. These self-administered formats enable clients to test for the presence of symptoms in a cost- and time-efficient manner in comparison to a clinical assessment. These forms are used in a variety of research and clinical contexts. This format appears to be more accessible and feasible to use in research studies than other formats. The limitation of report forms is that it is insufficient to diagnose adult ADHD. Clinicians incorporate self-reports or observer reports into a battery of assessments and consider it a valuable source of collateral and subjective experience.

The instruments referred to as diagnostic instruments were formatted as either semi-structured or structured interviews. These instruments require the clinicians to test for the diagnosis of adult ADHD by means of clinical interviews of patients. Even though these instruments provide for the diagnosis of adult ADHD, it is less frequently used. These measures generally require more time and is more costly than self-administered tools. The report forms are seen as an adjunct to a clinical interview with the client when making a diagnosis. However, information on the psychometric properties of these instruments is needed to conclude the appropriateness of the utilization of these tools.

#### **4.3.5 Domains of Instruments**

This section reports on the domains of the identified instruments and constructs thought to be associated with ADHD. In this section, information of the domains was divided between screening and diagnostic instruments. This was done to determine whether there are predictable

differences between the constructs measured by screening and diagnostic tools. A further distinction was made among the reporting on the screening instruments specific to adult ADHD and those reporting on multidimensional screening instruments. Table 4.9 summarises the domains itemized by the different adult ADHD specific screening instruments.

**Table 4.9 Domains of adult ADHD specific screening instruments**

<b>Instrument</b>	<b>Domains</b>
ADHD-SQ	Inattention Hyperactivity-Impulsivity
ASRS-v1.1	Inattentive Hyperactivity-Impulsivity Level of impairment
BAARS-IV	Inattention Sluggish Cognitive Tempo Hyperactivity-Impulsivity
Brown ADD scale	Organizing and activating to work Sustaining attention and concentration Sustaining energy and effort Managing affective interference Utilizing working memory and accessing recall
CAARS	Inattention/memory problems Hyperactivity/restlessness Impulsivity/emotional liability Problems with Self-concept Inattention symptoms Hyperactivity-Impulsivity Total ADHD symptoms ADHD index Inconsistency Index
CSS	Inattentive Hyperactivity-impulsivity
IVA+Plus	Impulsive and commission errors Attending and omission errors
WURS	Attention difficulties Hyperactivity/restlessness Temper Affective liability Emotional over-reactivity Disorganization Impulsivity

Eight adult ADHD specific screening instruments were included in Table 4.9, namely, the ADHD-SQ; ASRS-v1.1; BAARS-IV; Brown ADD scale; CAARS; CSS; IVA+Plus; and WURS. These instruments typically measure a number of different domains theorized to be part of ADHD as a construct. Items on inattention, hyperactivity and impulsivity were however consistently prioritized by the adult ADHD specific screening instruments. Three instruments, namely, the ADHD-SQ; ASRS-v1.1 (short version); and CSS only included items relative to these core symptoms. Thus, these short screeners only seek to gain information sufficient for screening for the presence of adult ADHD symptomology.

Lengthier instruments generally report more sections and items than shorter screening instruments. For instance, the CAARS-L and the WURS report 66 items and 61 items respectively. These are both instruments standardized for screening adult ADHD, although the latter scale retrospectively reports on childhood symptoms. All these sections and items are designed to elicit information deemed necessary for an accurate determination of the presence or absence of ADHD symptomology. In this way, longer measures prioritize accumulating more information on patients' symptom presentation across situations than shorter measures. The CAARS and WURS in this way includes domains on self-concept and emotional liability or overactivity. Shorter screeners do not elicit this type of information. This is a definite shortcoming. Especially since instruments that are to perform well in the assessment of adult ADHD should measure the core symptoms, inattention; hyperactivity and/or impulsivity across numerous domains. Table 4.10 summarises the domains itemized by the different multidimensional screening instruments.

**Table 4.10 Domains of the multidimensional screening instruments**

<b>Instrument</b>	<b>Domains</b>
IPDE-SQ	Work Interpersonal relations Affects Reality Testing Impulse control
SCL-90-R	Somatization Obsessive - compulsive Depression Anxiety Phobic anxiety Hostility Interpersonal sensitivity Paranoid ideation Psychoticism

Two instruments (IPDE-SQ and SCL-90-R) were included in Table 4.10. These instruments were not specific to measuring adult ADHD. Studies did not report whether these screeners included items on the core symptoms of ADHD directly. Instead, studies report domains relative to a broad range of constructs. Based on the reported domains it would be inappropriate to comment on the applicability for measuring adult ADHD of these multidimensional screening instruments. Hence the necessity in reporting on how these instruments performed psychometrically. Furthermore, Table 4.11 summarises the domains presented by the different adult ADHD specific diagnostic instruments.

**Table 4.11 Domains of adult ADHD specific diagnostic instruments**

<b>Instrument</b>	<b>Domains</b>
ASIA	Inattention Hyperactivity-Impulsivity
CAADID	Inattention symptoms. Hyperactive-Impulsive symptoms. Assesses impairment for the amalgam of Inattention and Hyperactive-Impulsivity.
WRAADDS	Attention difficulties Hyperactivity/restlessness Temper Affective lability Emotional over-reactivity Disorganisation Impulsivity

Three diagnostic instruments specific to adult ADHD were included in Table 4.11. These instruments were the ASIA; CAADID; and WRAADDS. The ASIA reportedly only includes items on the core symptoms, whereas the remaining diagnostic instruments included the core symptoms as well as additional symptoms. These instruments prioritized similar domains to the lengthier adult ADHD specific screeners, such as emotional lability or overactivity. The major distinction between these instruments is that the diagnostic tools are semi-structured instruments which allows for prompting beyond the core symptoms. Hence, adult ADHD manifests beyond the core symptoms. Further prominent symptoms are memory problems; emotional dysregulation and disorganization (Ramsay, 2017). To follow is Table 4.12 which lists the domains presented by the multidimensional diagnostic instruments.

**Table 4.12 Domains of multidimensional diagnostic instruments**

<b>Instrument</b>	<b>Domains</b>
PDI-4	-
PRISM	-
SCID-5	Based on the DSM-5 criteria Focuses on the symptoms over the past 6 months Emergence of childhood symptoms, cross-situational symptoms and impairments

Three multidimensional diagnostic instruments were included in Table 4.12. The reports on the respective domains were not included for the PDI-4 and PRISM. The SCID-5 is based on the DSM-V diagnostic criteria.

Overall, diagnostic instruments aim to carefully assess both the presence of the core symptoms and the level of impairment across various contexts. Instruments that seek to gain an in-depth description of the patient's symptom presentation includes numerous subsections and items. This is helpful since the diagnosis of ADHD in adults is complicated (Ramsay, 2017). It requires a careful assessment of the history of symptoms and impairment (Ramos-Quiroga et al, 2015). Intrinsic to this process is the identification of the age of onset, current symptoms and the presence of impairment in at least two domains (e.g. academic, family, work).

#### **4.4 Theoretical frameworks**

Studies have not reported on the theoretical frameworks undertaken in the respective studies.

#### **4.5 Theoretical and Operational Definitions**

Theoretical definitions were provided as part of the literature review of most articles. However, there was no attempts to explicitly link the definition of adult ADHD to the instruments or indicate how the instrument operationalized ADHD. Studies have instead stated that the instruments were based on particular diagnostic criteria rather than a particular definition. In this way, the accuracy of each instrument's ability to identify ADHD in adults was based on a particular criterion as opposed to how authors theoretically and operationally defined ADHD in adults.



## 4.6 Norms

Studies have not included information on the norm groups. Although some have included a brief description of initial studies in which the respective instruments were validated, the norm groups were not reported on.

## 4.7 Diagnostic and Statistical Manual (DSM)

As mentioned before, the diagnosing of ADHD in adults can be difficult. In fact, according to the Diagnostic and Statistical Manual, or DSM, the symptoms of the disorder have changed a number of times. In an attempt to accurately identify the disorder, instruments have often been formulated based on a particular version of the DSM and the criteria contained therein. Table 4.13 illustrates the relationship between the respective instruments and the versions of the relevant versions of the DSM.



**Table 4.13 Adherence to DSM**

<b>INSTRUMENT</b>	<b>DIAGNOSTIC MANUAL</b>	<b>Criterion A</b>	<b>Criterion B</b>	<b>Criterion C</b>	<b>Criterion D</b>	<b>Criterion E</b>
ADHD-SQ	DSM-IV	Items	–	–	–	–
ASIA	DSM-V	144 items	1 item	1 item	1 item	1 item
ASRS v1.1 (Long V)	DSM-IV	18 items	–	Factored into A	–	–
ASRS v1.1 (Short V)	DSM-IV	6 items	–	–	–	–
BAARS-IV	DSM-IV	18 items	–	–	–	–
Brown ADD scale	–	40 items	–	Factored into A	–	–
CAADID	DSM V	Number of items not specified	Number of items not specified	Number of items not specified	Number of items not specified	Number of items not specified
CAARS (Long V)	DSM-IV	66 items	Factored into final score	Items	–	–
CAARS (Short V)	DSM-IV	42 items	–	Factored into A	–	–
CSS	DSM-IV	18 items	–	–	–	–
IPDE-SQ	–	–	–	–	–	–
IVA+Plus	–	–	–	–	–	–
PRISM	DSM-IV	–	–	–	–	–
SCID-5	DSM-V	Number of items not specified	Number of items not specified	Number of items not specified	Number of items not specified	–
SCL-90-R	DSM-IV	–	–	–	–	–
PDI-4	Validated with DSM-IV tools	–	–	–	–	–
WRAADDs	Utah criteria	4 items	–	–	–	–
WURS	DSM-V	Factored into B	61 items	–	–	–

Included in Table 4.13 are the entire sixteen instruments identified, as well as the short and long version of the ASRS and CAARS. In total eighteen instruments are reported on here.



Fourteen of these instruments were reportedly based on the DSM and assessed certain criterion specified therein. The DSM-IV was the basis for ten instruments (ADHD-SQ; ASRS-v.1 short; ASRS-v.1 long; BAARS-IV; CAARS-short; CAARS-long; CSS; PRISM; SCL-90-R; and PDI-4). Four instruments (ASIA; CAADID; SCID-5; and WURS) were based on the DSM-V. It was not determined on which criteria, if any, the Brown ADD scale; IVA+Plus and IPDE-SQ are based on. However, the WRAADDS is reportedly based on the Utah criteria. This specific criterion was formulated prior to the publication of the DSM-III and is based on symptoms observed in adults as opposed to children (Marchant et al, 2013).

The current version of the DSM is the fifth edition. As mentioned before, the DSM-V was published in 2012 (APA, 2013). Thus, we note that the DSM-IV criteria for the diagnosis of ADHD in adults is still pervasively used in good quality research published after 2012, as evaluated in the present study. Subsequent to the DSM-IV there have been two versions namely the DSM-IV-TR and the DSM-V (APA, 2013). Thus, the lack of adoption of the DSM-V criteria in research, as evidenced by the lack of revision of instruments, is concerning (Walls et al, 2017). Among the changes included in the DSM-V was a decrease in the number of symptoms required for the diagnosis of adult ADHD. The age of onset was changed from 7 years to 12 and examples relative to adults were included with symptom illustrations.

### ***Criterion A***

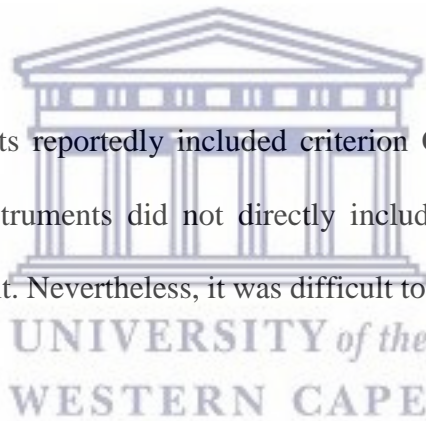
Twelve of the instruments contained items on criteria A of the ADHD. Most of the instruments had more items on criterion A than other criteria. The ASIA and the long version of the CAARS reported the 144 and 66 criterion A-inspired items respectively. This is indicative of instruments prioritizing the identification of the core symptoms associated with adult ADHD as stipulated in criterion A in both DSM-IV-TR and DSM-V.

### ***Criterion B***

Only four instruments included items based on the constructs included in criterion B. The WURS contains 61 items based on criterion B. However, this instrument is a retrospective self-report scale (Mórtbert et al, 2012). In completing this scale, patients report on their recollections of childhood symptoms across various domains. For instance; ‘As a child I was easily distracted’, or ‘As a child I was fidgety’ (Mórtbert et al, 2012). In this way, the WURS seeks to identify the presence of these symptoms during an adult’s childhood. Furthermore, other instruments, particularly self-report screeners have not prioritized retrospective reports of childhood symptoms. Perhaps because the focus of the screeners is to assess for the presence of symptoms at the time of assessment, rather than seek to meet all diagnostic requirements.

### ***Criterion C***

A number of instruments reportedly included criterion C-inspired items in the items related to criteria A. The instruments did not directly include items on criterion C, they reportedly accommodated for it. Nevertheless, it was difficult to determine the extent to which instruments accommodated it.



### ***Criterion D and E***

Items based on criteria D and E were not included in self-report and observer-report instruments. Criteria D and E require expertise and careful assessment which is difficult to capture in items on an instrument. It is better accomplished in a clinical interview covering both ADHD criteria and alternate explanations for symptoms.

Diagnostic instruments include items based on all criteria for ADHD. The ASIA, CAADID and the SCID-5 are semi-structured diagnostic interviews that include items on all criteria. These instruments are directly concerned with diagnosing adult ADHD and therefore ensures that all criteria are assessed. From this review the IPDE-SQ; PDI-4 and SCL-90-R are

multidimensional diagnostic interviews, but it could not be determined from the articles which criteria were assessed by these instruments.

It is considered best practice to base the diagnosis of ADHD in adults on the current DSM-V criteria. The DSM-V specifically addressed criticisms against previous DSMs, particularly the DSM-IV-TR. For example, the ADHD criteria contained in the DSM-IV-TR was not appropriate for measuring ADHD in adults (Hetchman, 2011). Thus, the continued use of criteria for ADHD derived from earlier iterations of the DSM is questionable and undesirable. The authors in the included articles did not report on the extent to which they engaged with the changes in diagnostic criteria for adult ADHD. The articles also did not make it clear if any of the items in the instruments were revised in an attempt to reflect the changes in diagnostic criteria in the DSM-V. Hence, research using instruments based on the DSM-V are considered more rigorous and valid than others based on outdated criteria.

#### **4.8 Psychometric Properties of Instruments**

The psychometric properties of the scales identified from good quality research were extracted and tabularized. Data was extracted for four psychometric properties, namely, Reliability, Validity, Sensitivity, and Specificity.

##### **4.8.1 Reliability**

Table 4.14 illustrates which articles reported on reliability for the respective instruments. It includes how the instruments performed in these particular studies. Studies that did not report on reliability were not included in Table 4.14.

**Table 4.14 Reliability**

Scale/(Author)	Test-Retest		Internal Consistency	Interrater reliability	
	CCC	ICC			r
ASRS v1.1 (Fuller-Killgore et al 2012)	-	-	0.876	-	
ASRS Korean V (Kim et al 2013)	0.878	-	0.885	-	
ASRS Turkish Long V (Evrn et al 2016)	0.765	-	0.863	-	
ASRS Turkish Short V (Evrn et al 2016)	0.636	-	0.654	-	
ASRS v1.1 Long V (Fuller-Killgore et al 2012)	-	-	0.876	-	
ASRS Short V (Fuller-Killgore et al 2012)	-	-	0.666	-	
ASIA (Takeda et al 2015)	-	-	0.64 – 0.92	0.97 – 1.00	
Brown scale (Fuller-Killgore et al 2012)	-	-	0.94	-	
CAARS:L (Amandor-Campos et al 2016)	High	-	High	Moderate	
CAARS:L (Fuller-Killgore et al 2012)	-	-	0.967	-	
CAARS:S (Amandor-Campos et al 2014)	High	-	High	Moderate	
CAARS-German (Christiansen et al 2012)	0.85 to 0.92	-	0.74-0.95	-	
CAARS (interrogated V) (Kooij et al 2013)	-	-	0.785-0.938	-	
ADHD-SQ (Manor et al 2012)	-	-	0.83; 0.80	-	
SCL-90-R (Eich et al 2012)	-	-	0.88	-	
WRAADDS (Marchant et al 2013)	-	0.96	0.78	0.75	

The table above summarizes the reliabilities reported by specific studies selected for the final extraction. The studies that reported on reliability generally included three forms of reliability: 1) Test-retest reliability; 2) Internal consistency; and Interrater reliability. However, as depicted in the table above, not all of the studies presented information across these forms.

**Test-retest reliability:** The table above indicates that only seven of the twenty-six studies selected have reported test-retest reliabilities. The study that communicated the highest test-retest correlation (0.96) was Marchant et al (2013) in their study on the psychometric properties of WRAADDS instrument. In this study, the test 1 and the test 2 were conducted with a two-week interval. Evrun et al (2016) reported the lowest test-retest reliability (0.636) for the short (6 items) Turkish Version of the ASRS-v1.1. The same study reported test-retest correlation for the long version (18 items) of 0.765 which was considered high. Amador-Campos et al

(2014); Amador-Campos et al (2016); Evrun et al (2016) and Takeda et al (2015) distinguished test-retest reliabilities of the different subscales within the specific instruments. Interestingly, the subscale, Problems with Self-Concept of the CAARS, illustrated lower correlations across versions.

The interval taken between test 1 and test 2 varied between studies. This suggests that there is no consensus of an optimal interval length for measuring the test-retest reliability for the constructs associated with ADHD. The study with the shortest interval (2 weeks) managed to achieve the highest correlation. It is possible that this high correlation was a result of transfer effects such as, the ability to recall items tested on in test 1. Other studies had greater intervals (4-6 weeks) yet reported satisfactory correlations.

**Internal consistency:** The internal consistency of instruments identified was reported in sixteen articles. Studies generally prioritised reporting on the internal consistency over the other reliabilities. Fuller-Killgore et al (2012) communicated the highest Cronbach alpha of 0.967, whereas Evrun et al (2016) reported the lowest (0.654) for the internal consistency of the short (6 items) ASRS-v1.1. Similarly, the short version of the ASRS-v1.1 presented a moderate (0.666) internal consistency in Fuller-Killgore et al (2012). Furthermore, other studies reported high internal consistencies in their studies. This is an indication of how the respective instruments performed in the studies, suggesting that the items of the instruments listed measured the proposed constructs in a consistent manner.

**Interrater reliability:** Only four studies reported on interrater reliability. The CAARS yielded moderate results in two studies (Amador-Campos et al, 2016; Amador-Campos et al, 2014). Other studies communicated high correlations signifying a greater stability between independent rater. The highest Cohen Kappa ( $k$ ) was reported by Takeda et al (2015) for the ASIA which ranged between 0.97 to 1.00.

#### 4.8.2 Validity

Table 4.15 below summarizes the types of validity reported by specific studies included in the final summation. Studies that did not report on validity were not included in table 4.15. To follow is a brief report on validity presented.

**Table 4.15 Validity**

Scale/(Author)	Concurrent	Discriminant
ASRS Korean V (Kim et al 2013)	high correlations	-
ASRS Turkish V (Evrin et al 2016)	-	( $t = -9.80, p < 0.001$ )
ASIA (Takeda et al 2015)	Acceptable	Acceptable
CAADID-Spanish (Ramos-Quiroga et al 2012)	Good	-
CAARS-German (Christiansen et al 2012)	moderately sig	highly sig
PDI-4 (Faries et al 2012)	0.75	$p < 0.05$
WRAADDs (Marchant et al 2013)	Good	$p < .001$

A total of six studies reported on validity of six different instruments. These studies were generally construction studies where validation was a stated objective. They commonly reported on: 1) Convergent validity; and 2) Discriminant validity.

**Convergent validity:** Few studies reported convergent validity. Kim et al (2013) presented the highest convergent validity, indicating that in that particular study, the Korean version of the ASRS-v1.1 was strongly correlated with the CAARS subscales. Other studies reported acceptable correlations. Christiansen et al (2012) reported that in their study the German version of the CAARS was moderately correlated with the BIS. Although correlations were highly significant between subscales, the correlation for the self-concept subscale was not significant. Additionally, the convergent validity of the German version of the CAARS and the WURS only reached significance for the impulsivity subscale in the same study.

**Discriminant validity:** Similar to convergent validity, few studies reported discriminant validity. Marchant et al (2013) reported a significant difference between the normative sample and those with ADHD for their report on the WRAADDs. This difference was consistent

across the seven domains: attention difficulties, hyperactivity/restlessness, temper, affective lability, emotional over-reactivity, disorganization, and impulsivity. This indicates that discriminant validity was significant ( $p < 0.001$ ) across all domains. Likewise, other studies that reported on this form of validity indicated significant statistical values.

**Interpretation of validity:** Overall studies reported acceptable validities. However, a substantially low number of studies reported on the conventional forms of validity. This has certainly changed from former studies associated with reporting on psychometric properties of instruments. One reason for this is that these studies often included instruments that were previously validated in other studies. Interestingly, the validation of instruments in these studies was conducted under specific objectives, namely; New instruments; New versions of instruments and instruments utilized in new conditions. Table 4.16 presents the studies that reported validity categorized by their objectives.

**Table 4.16 Validity in relation to study Objectives**

Validation of New Instruments	Properties of New Version of established instruments	Properties of the utilization of instruments in new conditions
Takeda et al (2015): ASIA	Christiansen et al (2012): German CAARS	Faries et al 2(012): PDI-4
	Ramos-Quiroga et al (2012): Spanish CAADID	Evren et al (2016): ASRS
	Evren et al (2016): Korean ASRS	Marchant et al (2013): WRAADDS

From the table above, it becomes apparent that three types of objectives or conditions occur in which researchers would specifically report on validity. First, validation studies of new instruments explicitly report on validity. Fewer of these types of studies were conducted and remain a focus of further research. Second, research using new versions of established tools tend to report on validity explicitly. In this instance the reporting of validity has to do with justifying the revisions made to established instruments. Third, studies using instruments in different or new conditions generally report validity explicitly. In this instance the focus is on

providing verification for the instrument with new target groups or contexts. In other words, the reporting of validity tends to occur when there is a measure of justification sought for new instruments, revised instruments and application to new contexts or with new target groups. In all of these studies, validation or, more broadly, psychometric properties, became an explicitly stated or implicit objective.

### 4.8.3 Sensitivity and Specificity

Table 4.17 summarizes the sensitivity and specificity reported by specific studies. Studies that did not report on sensitivity and specificity were not included in the table.

**Table: 4.17 Sensitivity and Specificity**

Scale/(Author)	Sensitivity	Specificity
ASRS v1.1 (Fuller-Killgore et al 2012)	7.10%	100.00%
ASRS v1.1 (Kingston et al 2013)	0.76	0.84
ASRS v1.1 (Söderström et al 2014)	90.2	35.0
ASRS v1.1 (van de Glind et al 2013)	0.88	0.67
ASRS Turkish V (Evrün et al 2016)	0.81	0.75
ASRS Screener (Daigre et al 2015)	86.7%	61.1%
ASRS Screener (Fuller-Killgore et al 2012)	21.40%	95.80%
BAARS-IV (Dvorsky et al 2016)	0.43-0.95	0.27-0.89
BAARS-IV (Young et al )	37.9	96.3
Brown scale (Fuller-Killgore et al 2012)	28.6	87.6
Brown scale (Kingston et al 2013)	0.84	0.73
CAADID-Spanish (Ramos-Quiroga et al 2012)	99%	67.68%
CAARS:L (Fuller-Killgore et al 2012)	7.10 – 50.00%	87.60 – 96.90%
CAARS:S (self report) (Kingston et al 2013)	0.63	0.91
CAARS:S (observer) (Kingston et al 2013)	0.76	0.75
CAARS-German (Christiansen et al 2012)	61.2-78.8	83.4-88
CSS (Söderström et al 2014)	85.4	40.0
ADHD-SQ (Manor et al 2012)	46%	95%
IPDE-SQ (Singh et al 2015)	84%	82%
PRISM (Ramos-Quiroga et al 2015)	90%	87.5%
SCL-90-R (Eich et al 2012)	75%	54%
PDI-4 (Faries et al 2012)	0.70	0.87
WURS (Daigre et al 2015)	79.6%	60.3%



A total of twenty-three articles included reports on these statistical measurements. To follow is a discussion on how the instruments performed regarding both sensitivity and specificity.

**Sensitivity:** Overall studies largely presented satisfactory sensitivities. Ramos-Quiroga et al (2012) reported the highest sensitivity. In this study, the Spanish version of the CAADID achieved 99% sensitivity. Instruments that achieved acceptable sensitivities displayed the ability to correctly identify ADHD within the respective studies. However, six studies reported figures below 50%, of which Fuller-Killgore et al (2012) reported the lowest sensitivity. This study presented unsatisfactory sensitivity for three self-report instruments including the ASRS-v1.1 (both the short and long (18 item) versions); the CAARS:L; and the Brown ADD. Authors argued that using self-reports to identify ADHD in a population was an inherent problem for establishing the sensitivity of the instruments. In doing so, the authors implied that the ability of self-report instruments to achieve acceptable sensitivity was compromised. However, these three instruments performed differently in other studies, suggesting greater variability in the sensitivity of these instruments in different contexts and with different target groups.

**Specificity:** To a large extent studies reported satisfactory specificity. Fuller-Killgore et al (2012) reported the highest specificity for the ASRS-v1.1. The ASRS achieved 100% in correctly identifying the non-ADHD population as not having ADHD. The other two instruments included in this study also indicated high specificity; CAARS:L (87.60 – 96.90%); and the Brown ADD (87.6%). Only three studies reported specificity indices below 50%. Söderström et al (2014) reported the lowest specificity for both instruments included within the study. In this study, the ASRS-v1.1 achieved 35.0% and the CSS 40%.

**Interpretation of Sensitivity and Specificity:** Substantially more studies have reported on Sensitivity and Specificity than other conventional types of validity. The statistical

calculation of Sensitivity and Specificity has become accepted practice in studies assessing the psychometric properties of instruments. In this way, authors were concerned with how accurately the instruments could discriminate between adults with and without ADHD rather than the general sense of validity. This is particularly true for instruments such as the CAARS; CAADID and ASRS that are considered to be well established at measuring ADHD in adults.

The reports indicate that there can be variability between sensitivity and specificity of instruments. For instance, Fuller-Killgore et al (2012) reported the lowest sensitivity, yet the highest specificity in their study on the psychometric properties of the ASRS-v1.1; Brown ADD; and CAARS;L three self-reports. In other studies, these three instruments yielded acceptable statistical figures for these indices. Despite the variances, overall, instruments performed well at identifying ADHD. The performance of instruments was largely dependent on the study.

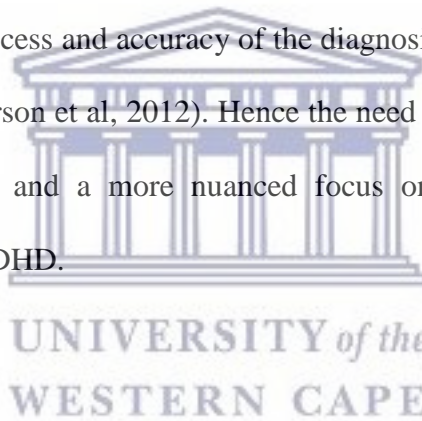


## CHAPTER FIVE: CONCLUSION

In this chapter an overview of the core findings is presented relative to the objectives of the study. This chapter concludes report on the present study by highlighting recommendations for further study, the significance of the study, and the limitations of the study.

### 5.1 Executive Summary

This study conducted a systematic review on studies reporting on instruments used in the assessment of adult ADHD. The aim was to consolidate evidence based on good quality literature for the purpose of validating the utilization of certain instruments over the use of others when assessing adult ADHD. Research repeatedly stressed the existence of countless uncertainties regarding the process and accuracy of the diagnosis of adult ADHD (Kooij et al, 2008; Liebenberg 2016; Asherson et al, 2012). Hence the need for filtering information using an extended search strategy, and a more nuanced focus on psychometric properties of instruments assessing adult ADHD.



### 5.2 Core Findings

This section presents the core findings in relation to review questions stipulated for this study. A total of twenty-six articles were recognized as good quality research based on the extended search employed in this study. Sixteen different instruments were identified from the articles included in the summation for assessing adult ADHD.

**Review question 1:** The first review question asked what instruments could be identified from good quality research to measure adult ADHD. Sixteen different instruments were identified for measuring ADHD in adults from the full texts included in the final summation. Eleven Adult ADHD specific instruments were identified including the ADHD-SQ; ASIA; ASRS; BAARS-IV; Brown ADD scale; CAADID; CAARS; CSS; IVA+plus;

WRAADDS; and WURS. Five multidimensional instruments were identified including the IPDE-SQ; PRISM; PDI-4; SCID-5; and SCL-90-R that are used to identify a broad range of disorders.

The instruments identified differed in terms of types, format, purpose, versions and language. The format and purpose of instruments were somewhat interrelated. For instance, screening instruments were generally formatted as self-reports. The self-reports merely identified the presence of symptoms of adult ADHD. Similarly, diagnostic instruments were formatted as clinician-based, semi-structured interviews. Diagnostic tools were clinician-based and were mainly associated with making informed diagnosis. This was consistent with the assertion by Ramsay (2017) that the diagnosis of ADHD required a comprehensive clinical interview covering both ADHD criteria and alternative explanations for symptoms.

The ASRS; CAARS; IPDE-SQ; PRISM; and SCID-5 reportedly all have different versions. Authors recognized the necessity of indicating which version they were using, as these diagnostic tools were intended to be reflective of the diagnostic criteria in the various versions of the diagnostic and statistical manual (DSM). Other reasons for having different versions in length is the time it takes to administer the screening. Shorter versions require less time but would then also provide less information. In this way, the identification of adult ADHD symptomology is based on limited information. It is then imperative that these short versions report acceptable sensitivity and specificity for the purpose of accurately identifying relevant symptoms. The gains made in time taken to administer must not be at the expense of robustness of the psychometric properties of the scale.

Many instruments were translated into different languages for use with specific population or target groups. Researchers have often sought to translate these instruments into other languages as opposed to developing new instruments relevant to specific languages and cultures. The articles included in the final summation included several examples of translated

instruments, and there was a patterned way in which the psychometric properties of instruments were reported when translation took place, or new versions of instruments were used. A notable criticism was that authors did not report on equivalency of measures. If equivalent forms exist that can be used across population or cultural groups, then many comparisons become possible. Thus, translated and equivalent versions are often desirable when internationalization and cross-cultural research agendas are at play.

**Review Question 2:** The second review question sought to identify the domains that are included in the measurement of the construct adult ADHD. Adult ADHD is characterized by poor functional outcomes such as employment difficulties, lower socioeconomic status, higher rates of divorce, more traffic violations and accidents, more criminality and incarcerations, more risky behaviour and high rates of comorbidity (Hechtman et al, 2011). The final summation identified two types of instruments, namely, screening instruments and diagnostic instruments. From the extracted data, a pattern was identified in the way domains were included in these instruments. Screening instruments typically focused on the core domains and included inattention, hyperactivity and impulsivity, as well as pervasiveness of symptoms across numerous contexts. Hechtman et al (2010) identified these domains as the core domains of adult ADHD that operationalize criterion A. For example, instruments such as the ADHD-SQ; CSS and ASRS-v1.1 only include domains related to these prominent symptoms. These instruments accentuated items directly associated with the core symptoms. Screening instruments include domains that are sufficient to determine the presence of core symptoms across contexts, but would not be able to specify the degree of impairment.

Diagnostic instruments typically included domains that corresponded with all criteria required for the diagnosis of adult ADHD. These instruments went beyond the core symptoms and included memory problems, emotional dysregulation, and disorganization as further prominent symptoms (e.g. ASRS, 2017). Instruments such as Brown ADD; CAARS;

WRAADDS and WURS include domains to accommodate these additional symptoms. In this way, such measures aim to carefully assess both the presence of the core symptoms, and the level of impairment, across various contexts. Thus, diagnostic instruments include domains that seek to gain an in-depth description of core symptoms as well as additional symptoms.

**Review question 3:** The third review question sought to establish the theoretical frameworks utilized in the different studies. Studies did not report on the theoretical frameworks, neither did they stipulate if it was completely omitted.

**Review question 4:** The fourth review question sought to establish which theoretical and operational definitions were used in measures of adult ADHD. The most striking finding from the extracted data was that the theoretical and operational definitions for the respective instruments were consistently not included or reported. In general, reference was made to the specific diagnostic criteria of adult ADHD and whether items relative to those criteria were included. The underlying assumption here was that the diagnostic and statistical manual provided criteria for diagnostic purposes, and that these were adopted in part of the whole as the default theoretical definition of instruments. The overall acceptance of the DSM as a diagnostic system meant that authors did not feel impressed upon to report explicitly on the theoretical definitions for adult ADHD as a syndrome. The findings related to theoretical definitions were intuitive findings given the status of the DSM. However, authors also did not report on operational definitions that were important to provide insight into how the theoretical constructs were actually measured by the items in the respective instruments. The accuracy of an instrument's measurement of adult ADHD is dependent on the correlation between items and particular diagnostic criteria. The emphasis in articles was on the universality and tacit agreement on the diagnostic criteria that make up the syndrome. This reality at a theoretical level is extended to operational definitions. Thus, there is a lack of reporting of theoretical and operational definitions for instruments measuring adult ADHD.

**Review question 5:** The fifth review question sought to know what reference or norm groups were used in measures of adult ADHD. Studies did not include information on the norm groups of the various instruments. The articles also did not identify source references that could provide comprehensive intelligence on information such as norm and reference groups.

**Review question 6:** The sixth review question asked how the respective measures reflected revisions of the diagnostic criteria in DSM-V. All the screening instruments were based on the DSM-IV criteria. Majority of the instruments reported adherence to the DSM-IV. The main critique of the former DSM-IV, ADHD criteria, was that the criteria failed in accurately identifying ADHD in adults (Hetchman, 2011).

Authors did not indicate whether they attempted to revise the items to approximate the DSM-V criteria for adult ADHD. The authors also did not identify the use of DSM-IV criteria as a limitation in their respective studies. A potential limitation of the present study is that it included studies between 2012 and 2016. Perhaps studies published later have documented such revisions. However, the revisions were brought about in 2012, and research into adult ADHD has lagged in the adoption of DSM-V criteria. The adaptation of existing and widely used instruments measuring adult ADHD to reflect current DSM-V criteria remain a focus of further research.

**Review question 7:** The seventh review question asked about the psychometric properties of measures of adult ADHD extracted from the final summation. This question dealt with reliability estimates

- **Reliability estimates**

Overall, satisfactory reliabilities were reported for the instruments. The short version of the ASRS-v1.1 scored lower for both test-retest and internal consistency than the longer form in the same study. The results indicate that in this specific study the longer ASRS-v1.1 instrument

was more reliable than the shorter form. In this way, even though the shorter is more accessible, time efficient, and more predictive of ADHD symptoms, the longer version reports more consistent reliability across items and tests. The CAARS reported the highest reliability across studies. This is consistent with the findings presented by Taylor et al (2011).

With regards to the test-retest reliability, the interval taken between test 1 and test 2 varied between studies. This suggests that there is a lack of agreement on what constitutes an optimal interval length for measuring the test-retest reliability. The study with the shortest interval (2 weeks) managed to achieve the highest test-retest correlation. It is possible that this high correlation was a result of transfer effects such as the ability to recall items tested in the first test. Other studies had larger intervals (4-6 weeks) but also reported satisfactory correlations.

- **Validity indicators**

Two trends were observed in relation to the reporting of validity indicators. First, fewer articles reported on the conventional forms of validity. The second observation was that there was a pattern to the reporting of validity indicators. Studies dedicated to validation as an objective tended to report validity indicators in detail. Three types of studies were identified, namely, validation studies of new instruments; validation of new versions of existing instruments and validity indicators of instruments used in new conditions. Thus, we observe that more detailed reporting on validity was associated with validation studies rather than a characteristic of all research. The lack of consistent reporting of validity indicators specifies that researchers defer to earlier validation of instruments, and prioritize reporting of research outcomes above reporting on psychometric properties of instruments. This is a concerning trend in that there is no engagement in the reporting with the robustness of instrumentation that produces the research results.



- **Sensitivity and specificity estimates**

Two trends were observed in relation to sensitivity and specificity. First, the validation studies reported on specificity and sensitivity over and above conventional validity indicators. Second, of the studies that did not specifically set out to examine validation, twelve articles reported on the sensitivity and specificity of their instruments but did not report on conventional validity indicators. The statistical calculation of sensitivity and specificity has become popular and is increasingly reported. The extracted data here suggested that authors select instruments that are reliable and valid. Thus, these conventional psychometric properties are considered to be implicit in the methodology of studies. Authors demonstrate their concern about the accuracy of the instruments in identifying ADHD in adults by focusing on the sensitivity and specificity of instruments. Authors demonstrate their commitment to this by explicitly reporting on specificity and sensitivity.

The results indicate that specificity and sensitivity are interrelated. Researchers must track and examine the relationship between specificity and sensitivity in the context of their respective studies. In particular, sensitivity scores must be interpreted accurately. Akobeng (2006) stated that sensitivity only communicates how good an instrument is. Thus, researchers should take care not to over-interpret low sensitivity scores if studies do not include positively diagnosed ADHD populations. It is difficult to evaluate how well instruments can correctly identify ADHD if the sample did not include positively diagnosed individuals.

### **5.3 Conclusion**

There is a clear body of literature reporting on instruments measuring adult ADHD. Through this filtration process 16 instruments were identified that have been used to assess adult ADHD. Two types of instruments were identified namely screening tools and diagnostic instruments. Screening tools assessed for the presence of core symptoms reflected in Criteria A of the diagnostic and statistical manuals and use largely self-report. Diagnostic instruments were

multidimensional and clinician-based. The two types of instruments correspond to two distinct processes: screening and diagnosis. Screens identify the presence of symptoms, but do not provide any indication of how these symptoms relate to differential diagnoses. Diagnosis requires clinician assessment that clearly differentiates the syndrome from symptoms and determines whether it is at disorder level i.e. the extent of impairment.

The final summation revealed that the instruments were based on the DSM IV criteria that has been criticized in the literature as insufficient for the assessment of ADHD. There was no explicit indication given that items were revised to more accurately reflect DSM-V criteria for the syndrome. The ASIA; CAADID; SCID-5 were the only diagnostic instruments that were DSM-V criteria. The Structured Clinical Interview Disc (SCID) is used in research settings specifically to conduct research that is used to inform revisions of the DSM and thus the instrument has various forms that correspond to the respective versions of the DSM. Therefore, the SCID-V is specifically based on the DSM-V. Hence, the need to revise well-established instruments with high levels of adoption to more accurately reflect DSM-V criteria for adult ADHD remains a focus of further research.

The final summation revealed that there are clear patterns to how psychometric properties were reported. Conventional forms of validity are assumed to be sufficiently dealt with in published validation literature. The decision to use any given instrument is assumed to be an informed decision and sufficient evidence of the basic psychometric properties of the instrument. Psychometric properties are more likely to be reported explicitly in validation studies. Researchers increasingly use sensitivity and specificity to provide an indication of how accurate instruments were in discerning the presence or absence of disorder (adult ADHD). A cautionary note is that these indices must be interpreted carefully. For example, the findings illustrated how the sensitivity index is adversely impacted when samples do not include

positively diagnosed individuals. The instruments identified all reported satisfactory psychometric properties.

## 5.4 Limitations

The following limitations were identified for the present study:

- The selection of the databases available at UWC was done through a very thorough process. However, the strength of the search strategy was limited by the selection of databases that the university subscribed to. The decision to not access complimentary databases through other facilities, such as neighbouring universities, was pragmatic but limiting, nonetheless. The study did not explore the pattern of subscription to databases and thus bias might inadvertently have been introduced into the study.
- The decision to exclude articles published in foreign languages introduced language bias.
- The decision to exclude articles that required payment in order to access introduced bias. This decision was based on what the general university community would have access to. Eight articles were identified in the comprehensive search of UWC databases but were not available as full text without payment. Excluding these articles on the basis of costs rather than methodological rigour or relevance was a potential source of bias and therefore a limitation. Failure to explore other avenues such as inter-library loan or seeking out membership at the library of sister universities were reasonable solutions that were not considered at the time. This was a limitation of the present study.
- The critical appraisal tool was very detailed and more comprehensive than other tools available. Given the rigour of the tool, the threshold score could have been set lower. Thus, good quality articles might have been excluded from the final summation due to the overly stringent criterion set for the full text review.

## 5.5 Recommendation for future study

- It is recommended that the study be replicated with revised inclusion/ exclusion criteria such as the inclusion of foreign language articles. The proviso would be that foreign languages with strong academic traditions be considered for inclusion such as French, Spanish and Flemish/ Dutch.
- It is recommended that the study be replicated with an expanded search strategy that includes medical databases with publications from psychiatry, psychiatric nursing, consultation liaison, and forensic applications.
- It is recommended that the study be replicated, and the threshold score be set lower given the comprehensiveness of the SFS scoring system.
- From the findings, systematic investigation of well-established instruments is recommended, with an explicit focus on the adaptations required to more accurately reflect DSM-V criteria.
- It is recommended that health professionals and researchers into ADHD be studied in order to gain insight into perceptions about the revisions of the adult ADHD criteria, as well as the adoption of DSM-V criteria for the syndrome.
- It is further recommended that future studies examine the perceptions about the reasons underlying the lack of research into revisions of instruments to reflect DSM-V criteria.
- It is recommended that the level of understanding and adoption of sensitivity and specificity of instruments in clinical practice and research on adult ADHD be promoted.

## 5.6 Significance of the study

The present study made several important contributions.

1. The study demonstrated the need for filtration in the problem formulation. This was important as many systematic reviews do not demonstrate this level of application in the conceptualization.

2. The study adopted recommended best practice guidelines in the planning and execution of the systematic review methodology, such as multiple reviewers. This enhanced the methodological rigour and coherence of the study. In this way the study provides a good template for good quality systematic review.
3. The study used the PRISMA to guide the operational steps of the study, as well as the reporting. The use of this internationally recognized flowchart and reporting system meant that the study adhered to internationally recognized practices that are preferred in systematic review methodology. This meant that the study conformed to best practices and provided a good template for other student researchers to consider.
4. The study used a more comprehensive critical appraisal tool that provided a more robust operationalization of methodological rigour and quality.
5. The study used three levels of reporting that enhanced the quality of the meta-synthesis beyond a purely descriptive level. Thus, the study made important contributions at the level of methodology and the practice of secondary research.
6. The study consolidated the literature reporting on instruments measuring adult ADHD. Through this filtration process, the identification of instruments produced a summary of available tools and identified patterns of reporting biases towards research outputs at the expense of instrumentation analysis, especially psychometric properties. The study identified a deference to clinical references at the expense of good research practice. The study also identified that there is a need to systematically examine underlying issues in the lack of adoption of DSM-V criteria in research on adult ADHD. In this way, the study made a practical contribution to both practice and research by highlighting these gaps and biased practices.
7. The study identified that there is a need for caution in clinical practice when selecting instruments for the assessment of adult ADHD.

## Reference list:

- Alder, L. A., Solanto, M., Escobar, R., Lipsuis, S., & Upadhyaya, H. (2016). Executive functioning outcomes over 6 months of atomoxetine for adult ADHD: Relationship to maintenance of response and relapse over the subsequent 6 months after treatment. *Journal of Attention Disorder*, 1(10).
- Alderson, R. M., Hudec, K. L., Patros, C. H. G., & Kasper, J. (2013). Working memory deficits in adults with Attention-Deficit/Hyperactivity Disorder (ADHD): An Examination of central executive and storage/rehearsal processes. *Journal of Abnormal Psychology*, 122(2), 532-541.
- Amador-Campos, J.A., Gómez-Benito, J., Ramos-Quiroga, J. A. (2014). The Conners adult ADHD rating scale-short self-report and observer forms: Psychometric properties of the Catalan version. *Journal of Attention Disorders*, 18(8), 671-679.
- Amador-Campos, J.A., Nuno, L., Gómez-Benito, J. (2016). The Conner's adult ADHD rating scales-long self-report and observer forms: validation of the Catalan version. *Journal of Psychopathological Behaviour Assessment*, 38, 209-2015.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5<sup>th</sup>ed.). Arlington, VA: American Psychiatric Publishing.
- Asherson, P., Akehurst, R., Kooij, J. J. S., Huss, M., Beusterien, K., Susané, R., Gholizadeh, S., & Hodgkins, P. (2012). Under-diagnosis of adult ADHD: Cultural influences and societal burden. *Journal of Attention Disorders*, 16(5), 205-385.
- Atwoli, L., Owiti, P., & Ndambuki, D. (2011). Attention deficit hyperactivity disorder symptom report among medical students in Elboret, Kenya. *African Journal of Psychiatry*, 286-289.

- Barkley, R. A., & Murphey, K. R. (2010). Impairment in occupational functioning and adult ADHD: The predictive utility of executive function (EF) ratings versus EF tests. *Archives of Clinical Neuropsychology*, 25, 157-173.
- Brod, M., Johnston, J., Able, S., Swindle, R. (2006). Validation of the adult attention deficit hyperactivity disorder quality of life scale (AAQoL): A disease specific quality of life measure. *Quality of life research*, 15(1), 117-129.
- Brusilovsky, P., Ahn, J., & Rasmussen, E. (2010). Teaching information retrieval with web based interactive visualisation. *Journal of Education for Library and Information Science*, 51(3), 187-200.
- Christiansen, H., Kis, B., Hirsch, O., Matthies, J., Hebebrand, J., Uekermann, M., Abdel-Hamid, M., Kraemer, M., Wiltfang, J., Graf, E., Colla, M., Sobanski, E., Alm, B., Rösler, M., Jacob, C., Jans. T., Huss, M., Schimmelmann, B. G., & Philipsen, A. (2012). *European Psychiatry*, 27, 321-328.
- Daigre, C., Roncero, C., Rodriguez-Cintas, L., Ortega, L., Lligona, A., Fuentes, S., Perez-Pazos, J., Martinez-Luna, N., & Casas, M. (2010). Adult ADHD screening in alcohol dependent patients using the Wender-Utah rating scale and the adult ADHD self-report scale. *Journal of Attention Disorders*, 19(4) 328-334.
- De Quiros, G., & Kinsbourne, M. (1998). Adult ADHD: Analysis of self-ratings on a behaviour questionnaire. *Adult ADHD*, 1(1), 140-147.
- Dvorsky, M. R., Langberg, J. M., Molitor, S. J., & Bourchtein, E. (2016). Clinical utility and predictive validity of parent and college student symptom ratings in predictive ADHD diagnosis. *Journal of Clinical Psychology*, 72(4), 401-418.

- Eich, D., Angst, J., Frei, A., Ajdacic-Gross, V., Rössler, W., & Gamma, A. (2012). A new rating scale for adult ADHD based symptom checklist 90 (SCL-90-R). *European Arch Psychiatry Clinical Neuroscience*, 262, 519-528.
- Epstein, J. N., & Kollins, S. H. (2006). Psychometric properties of an adult ADHD diagnostic interview. *Journal of Attention Disorders*, 9(5), 504- 514.
- Erhardt, D., Epstein, J. N., Conners, C. K., Parker, D. A., & Sitarenios, G. (1999). Self-ratings of ADHD symptoms in adults II: Reliability, validity, and diagnostic sensitivity. *Journal of Attention Disorders*, 3(3), 153-158.
- Evren, C., Umut, G., Bozkurt, M., Teksin-Unal, G., Agachanli, R., Evren, B. (2016). Psychometric properties of the Turkish version of the adult ADHD self-report scale (ASRS-v1.1) in a sample of inpatients with alcohol use disorders. *The Journal of Psychiatry and Neurological Sciences*, 29, 109-119.
- Faries, D. E., Houston, J. P., Sulc, E. M., & Swindle, R. W. (2012). A cross-validation of the provisional diagnostic instrument (PDI-4). *BMC Family Practice*, 13, 104.
- Foxcroft, C. D. (2011). Ethical Issues Related to Psychological Testing in Africa: What I Have Learned (So Far). *Online Readings in Psychology and Culture*, 2(2). <https://doi.org/10.9707/2307-0919.1022>.
- Fuller-Killgore, M. D., Burlison, J., & Dwyer, W. (2012). Comparison of three ADHD screening instruments in college students of varying cognitive ability. *Journal of Attention Disorders*, 17(5), 449-454.



- Giacobini, M., Medin, E., Ahnemark, E., Russo, L. J., Carlquist, P. (2012). Prevalence, patient characteristics, and pharmacological treatment of children, adolescents and adults diagnosed with ADHD in Sweden, 1-11.
- Gjervan, B., Hjemdal, O., & Nordahl, H. M. (2016). Functional impairment mediates the relationship between adult ADHD inattentiveness and occupational outcome. *Journal of Attention Disorders*, 20(6), 510-518.
- Gjervan, B., Torgersen, T., Nordahl, H. M., & Rasmussen, K. (2012). Functional impairment and occupational outcome in adults with ADHD. *Journal of Attention Disorders*, 16(7), 544-552.
- Godfrey, C. M., & Harrison, M. B. (2012). *Can synthesis is a quick reference resource to guide the use of the Joanna Briggs Institute methodology of synthesis*. Queens; Joanna Briggs Collaboration version 3.0.
- Gorlin, E. I., Dalrymple, K., Chelminski, I., & Zimmerman, M. (2016). Reliability and validity of a semi-structured DSM-based diagnostic interview module for the assessment of attention deficit hyperactivity disorder in adult psychiatric outpatients. *Psychiatric Research*, 242, 46-53.
- Gray, S., Woltering, S., Mawjee, K., & Tannock, R. (2013). The adult ADHD self-report scale (ASRS): utility in college students with attention-deficit/hyperactivity disorder. *PeerJ* 2:e324; DOI 10.7717/peerj.324.
- Grinnell, M. (2011). *Adult ADHD in DSM-IV-TR and DSM-5: impact of increased age of onset on prevalence*. University of North Carolina.
- Halperin, J. M., Bédard, A. V., & Curchach-Lichtin, J. T. (2012). Preventative Interventions for ADHD: A neurodevelopmental perspective. *Neurotherapeutics*, 9, 531-541.

- Hatchman, L. (2011). Prospective follow-up studies of ADHD: Helping establish a valid diagnosis in adults. *Journal of the American academy of child & adolescents psychiatry*, 50(6), 533-535.
- Hechtman, L., French, L. R., Mongia, M., Cherkasova, M. V. (2011). Diagnosing ADHD in adults: limitations to DSM-IV and DSM-V proposals and challenges ahead. *Neuropsychiatry*, 1(6), 579–590
- Hesson, J., & Fowler, K. (2015). Prevalence and correlates of self-reported ADD/ADHD in large national sample of Canadian adults. *Journal of Attention Disorders*, 1-10.
- Higgins, J. P., & Green, S. (2006). *Cochrane handbook for systematic reviews of interventions* 4.2.6. Chichester, UK: John Wiley & Sons.
- Kessler, R. C., Adler, L. A., Gruber, M. J., Sarawate, C. A., Spencer, T., & Van Brunt, D. L. (2007). Validity of the World Health Organisation adult ADHD self-report scale (ASRS) screener in a representative sample of health plan members. *International Journal of Methods in Psychiatry Research*, 16(2), 52-65.
- Kim, J., Lee, E., & Joung, Y. (2013). The WHO adult ADHD self-report scale: Reliability and validity of the Korean version. *Psychiatry Investic*, 10, 41-46.
- Kingston, D. A., Ahmed, A. G., Gray, J., Bradford, J., & Seto, M. C. (2013). The assessment and diagnosis of attention deficit hyperactivity disorder in adult forensic psychiatric outpatients. *Journal of Psychopathological Behaviour Assessment*, 35, 293-300.
- Kooij, S. J., Boonstra, M. A., Swinkels, S. H. N., Bekker, E. V., De Noord, I., & Buitelaar, J. K. (2008). Reliability, validity and utility of instruments for self-report and informant report concerning symptoms of ADHD in adult patients. *Journal of Attention Disorders*, 11(4), 445-458.

- Kooij, S. J., Conners, K. C., Goto, T., Tanaka, Y., Williams, D., Allen, A. J., Heinloth, A. N., & Upadhyaya, H. P. (2013). Validity of Conners adult attention deficit hyperactivity/impulsivity disorder rating scale investigator rated: screening version in a patients from within and outside Europe, *Psychiatry Research*, 208, 94-96.
- Laher, S., & Cockcroft, K. (2014). Psychological assessment in post-apartheid South Africa: The way forward. *South African Journal of Psychology*, 44(3), 303-314.
- Liebenberg, R. (2016). Challenges of diagnosing adult ADHD. *Mental Health Issues*, 20-22.
- Manor, I., Vurembrandt, N., Rozen, S., Gevah, D., Weizman, A., & Zalsman, G. (2012). Low self-awareness of ADHD in adults using a self-report screening questionnaire. *European Psychiatry*, 27, 314-320.
- Marchant, B. K., Reimherr, F. W., Robison, D., Robison, R. J., & Wender, P. H. (2013). *Psychological Assessment*, 25(3), 942-950.
- Marshall, P., Hoelzle, J. B., Heyerdahl, D., & Nelson, N. W. (2016). The impact of failing to identify suspect effort in patients undergoing adult attention-deficit/hyperactivity disorder (ADHD) assessment. *Psychological Assessment*, 28(10), 1290-1302.
- McCann, B. S., & Roy-Byrne, P. (2004). Screening and diagnostic utility of self-report attention deficit hyperactivity disorder scales in adults. *Comprehensive Psychiatry*, 45(3), 175-183.
- McCann, B. S., Scheele, L., Ward, N., & Roy-Byrne, P. (2000). Discriminant validity of Wender Utah rating scale for attention deficit hyperactivity disorder in adults. *Neuropsychiatry Clinical Neuroscience*, 12(2), 240-245.

- McGough, J. J., & Barkley, R. A. (2004). Diagnostic controversies in adult attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 161, 1948-1956.
- McMillan, J. H., & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5<sup>th</sup> ed.). New York: Longman.
- Mehringer, A. M., Downey, K. K., Schuh, L. M., Pomerleau, C. S., Snedecor, S. M., & Schubiner, H. (2002). The Assessment of hyperactivity and attention: Development and preliminary validation of a brief self assessment of adult ADHD. *Journal of Attention Disorders*, 5(4), 223-231.
- Mörtberg, E., Tilfors, K., & Bejereot, S. (2012). Screening for ADHD in an adult social phobia sample. *Journal of Attention Disorders*, 16(8), 645-649.
- Munnik, E. & Smith, M. R. A systematic review of screening tools for social/emotional competence as a domain of school readiness. In Smith, M.R. (Chair). *Symposium on Methodological Rigour and Coherence: Deconstructing The Quality Appraisal Tool In Systematic Review Methodology* conducted at the 21<sup>st</sup> National Conference of the Psychological Association of South Africa, South Africa, 2015.
- Nelson, J. M., & Gregg, N. (2012). Depression and anxiety among transitioning adolescents and college students with ADHD, dyslexia or comorbid ADHD/dyslexia. *Journal of Attention Disorders*, 16(3), 244-254.
- Prevatt, M., & Young, J. L. (2014). Recognising and treating Attention deficit/Hyperactivity Disorder in college students. *Journal of College Student Psychotherapy*, 28, 182-200.
- Price, M., & Raffelsbauer, D. (2012). ADHD: A true neurodevelopmental disorder? *Medical Writing Journal*, 21(2), 114-122.

- Prosser, B., & Reid, R. (2013). The DSM-5 changes and ADHD: More than a tweak of terms. *Australian and New Zealand Journal of Psychiatry*, 47(12), 1196-1203.
- Ramos-Quiroga, J. A., Bosch, R., Richarte, V., Valero, S., Gómez-Barros, N., Nogueira, M., Palomar, G., Corrales, M., Sáez-Francás., Corominas, M., Real, A., Vidal, R., Chalita, P. J., & Casas, M. (2012). Criterion and concurrent validity of Conners adult ADHD diagnostic interview for the DSM-IV (CAADID) Spanish version. *Rev Psiquiátr Saud Mental*, 5(4), 229-235.
- Ramos-Quiroga, J. A., Diaz-Digon, L., Comin, M., Bosch, R., Palomar, G., Chalita, J. P., Roncero, C., Nogueira, M., Torrens, M., & Miguel, C. (2015). Criteria and concurrent validity of ADHD section of the psychiatric research interview for substance and mental disorders. *Journal of Attention Disorders*, 19(12), 999-1006.
- Ramsay, J. R. (2016). "Turning intentions into actions": CBT for adult ADHD focused on implementation. *Clinical Case Studies*, 15(3), 179-197.
- Ramsay, J. R. (2017). Assessment and monitoring of treatment response in adult ADHD patients: current perspectives. *Neuropsychiatric Disease and Treatment* (13), 221–232.
- Rösler, M., Retz, W., Thome, J., Schneider, M., Stieglitz, R. D., Falkai, P. (2006). Pathological rating scales for diagnostic use in adults with attention deficit hyperactivity disorder (ADHD). *Europe Psychiatric Clinical Neuroscience*, 256(1), 3-11.
- Sadock, B. J., Sadock, V. A., & Ruiz, P. (2015). *Kaplan & Sadock's Synopsis of Psychiatry* (11<sup>th</sup>ed.). Philadelphia: Wolters Kluwer.
- Sanchez-Gacia, M., Fernandez-Calderon, F., Carmona-Marquez, J., Chico-Garcia, M., Velez-Moreno, A., & Perez-Gomez, L. (2015). Psychometric properties and adaption of the

ASRS in a Spanish sample of patients with substance use disorders, *Psychological Assessment*.

Sandelowski, M., Docherly, S., & Emden, C. (1997). Qualitative meta-synthesis: Issues and techniques. *Research in Nursing and Health*, 20, 365-371.

Sibley, M. H., Pelham, W. E., Molina, B. S. G., Gnagy, E. M., Waxmonsky, J. G., Waschbusch, D. A., Derefinko, K. J., Wymbs, B. T., Garefino, A. C., Babinski, D. E., & Kuriyan, A. B. (2012). When diagnosing ADHD in young adults emphasize important reports, DSM items and impairment. *Journal of Consulting and Clinical Psychology*, 80(6), 1052-1061.

Singh, P., White, S., Saleem, K., & Luty, J. (2015). Identifying ADHD in adults using the international personality disorder examination screening questionnaire. *Journal of Mental Health*, 24(4), 236-241.

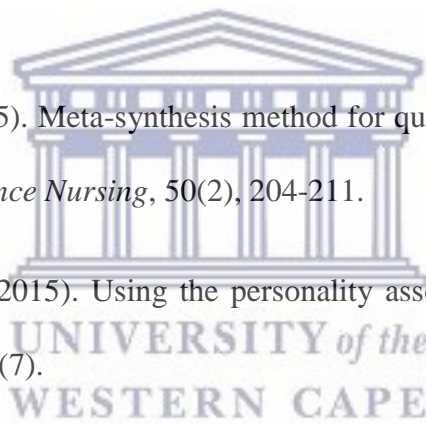
Smith, M.R. Methodological Rigour and Coherence: A concept paper. In Smith, M.R. (Chair). *Symposium on Methodological Rigour And Coherence: Deconstructing The Quality Appraisal Tool In Systematic Review Methodology* conducted at the 21<sup>st</sup> National Conference of the Psychological Association of South Africa, South Africa, 2015.

Smith, M.R., Franciscus, G. Swartbooi, C. Munnik, E. & Jacobs W. The SFS scoring system. In Smith, M.R. (Chair). *Symposium on Methodological Rigour And Coherence: Deconstructing The Quality Appraisal Tool In Systematic Review Methodology* conducted at the 21<sup>st</sup> National Conference of the Psychological Association of South Africa, South Africa, 2015.

Söndestrom, S., Pettersson, R., & Nilsson, K.W. (2014). Quantitative and subjective behavioural aspects in the assessment of attention-deficit hyperactivity disorder (ADHD) in adults. *Nord J Psychiatry*, 68, 30-37.

- Spencer, T. J., Adler, L. A., Qiao, M., Saylor, K. E., Brown, T. E., Holdnack, J. A., Schuh, K. J., Trzepacz, P. T., & Kelsey, D. K. (2009). Validation of the adult ADHD investigator symptom rating scale (AISRS). *Journal of Attention Disorders*, 14(1), 57-68.
- Sprafkin, J., Gadow, K. D., Weiss, M. D & Schneider, J. (2007). Psychiatric comorbidity in ADHD symptoms subtypes in clinic and community adults. *Journal Attention Disorders*, 11(2), 114-124.
- Takeda, T., Tsuji, Y., Uwatoko, T., & Kurita, H. (2015). Reliability and validity of ADHD diagnostic criteria in the Assessment System for individuals with ADHD (ASIA): a Japanese semi-structured diagnostic interview, *BMC Psychiatry*, 15(130).
- Taylor, A., Deb, S., & Unwin, G. (2011). Scales for the identification of adults with attention deficit hyperactivity disorder (ADHD). *Research in Developmental Disabilities*, 32(1), 924-938.
- Teing, L. S. (2007). Systematic Review made simple for nurses. *SGH Proceedings*, 16(2), 104-110.
- Van der Glind, G., van den Brink, W., Koeter, W. J., Carpentier, P., Kaye, S., Bu, E. H., Skutle, A., Frank, J., Konstenius, M., Moggi, F., Dom, G., Verspreet, S., Demetrovics, Z., Kapitany-Foveny, M. K. (2013). Validity of the adult ADHD self-report scale (ASRS) as a screener for adult ADHD in treatment seeking substances use disorder patients. *Drug and Alcohol Dependence*, 132, 587-596.
- Van der Westhuizen, A. (2010). Attention deficit hyperactivity disorder (ADHD). *SA Pharmaceutical Journal*, 10-20.

- Van Schalkwyk, G. L., & Schronen, J. (2011). Attention deficit/hyperactivity disorder in adults: a guide for the general practitioner. *South Africa Family Practitioners*, 53(2), 131-134.
- Vidal, R., Valero, S., Nogueira, M., Palomar, G., Corrales, M., Richarte, V., Bosch, R, Gómez-Barros, N., Corominas, M., Casas, M., & Ramos-Quiroga, A. J. *Comprehensive Psychiatry*, 55, 1712-1719.
- Walls, B. D., Wallace, E. R., Brothers, S. L., Berry, D. T. R. (2017). Utility of the Conners' Adult ADHD Rating Scale Validity Scales in Identifying Simulated Attention-Deficit Hyperactivity Disorder and Random Responding. *Psychological Assessment*, 29(12), 1437 – 1446.
- Walsh, D., & Downe, S. (2005). Meta-synthesis method for qualitative research: A literature review. *Journal of Advance Nursing*, 50(2), 204-211.
- Watson, J., & Liljequist, L. (2015). Using the personality assessment inventory to identify ADHD-like symptoms.1(7).
- Weiss, M. D., & Weiss, J. R. (2004). A guide to the treatment of adults with ADHD. *J clin psychiatry*, 65(3), 27-37.
- Weyandt, L. L., & DuPaul, G. J. (2012). Introduction to a special series on college students with ADHD: Psychological issues, comorbidity and treatment, *Journal of Attention Disorders*, 16(3), 199-201.
- Wilens, T. E., Faraone, S. V., & Biederman, J. (2004). Attention deficit hyperactivity disorder in adults. *American Medical Association*, 292(5), 619-623.





Yeh, C. B., Gau, S. S., Kessler, R., & Wu, Y. (2008). Psychometric properties of the Chinese version of the adult ADHD self-report scale. *International Journal of Methods in Psychiatric Research*, 17(1), 45-54.

Young, S., González, R. A., Mutch, L., Lallet-Lambert, I., O'Rourke, L., Hickey, N., Asherson, P., & Gudjonsson, G. H. (2016). *Psychological Medicine*, 46, 1449-1458.







## Appendix C: Critical Appraisal Tool

### CRITICAL APPRAISAL TOOL

#### SYSTEMATIC REVIEW


Author: \_\_\_\_\_

Title: \_\_\_\_\_

Source: \_\_\_\_\_

#### Section A: Methodological Rigour (Maximum score = 44).

##### Purpose:

- 
- 1) Is a clear problem statement present?
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 2) Is a clear statement of the aims of the research made? (Goal, relevance, why it was thought to be important)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 3) Was recent literature accessed and discussed?
- |           |           |                 |
|-----------|-----------|-----------------|
| >10 years | 6-10years | 1-5 year period |
| 1         | 2         | 3               |

##### Design:

- 4) Is the design of the study identified and described in detail? (How the study will be conducted)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 5) Is the research design appropriate to address the aims of the research? (Did the researchers discuss why they decided to use *quantitative*- RCT, Cohort, Single case, case study, cross sectional etc. OR *qualitative* - focus group, semi-structured interview etc. OR a *mixed method* research design?)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |

##### Sample:

6) Is the sample described in detail?  

No	Yes
1	2

7) Was the recruitment strategy appropriate for the aims of the research?  

No	Yes
1	2

Is it a probability or non-probability sample?  

Not mentioned	Probability	Non- probability
1	2	3

**Hint:** If it is non-probability sampling, did they test the sample to see if it fits the criteria. Did they report on it? (Then it qualifies for 2)

8) Is the sample size greater than 50? If not, is a formula computed to help with sample size?  

No	Yes
1	2

**Hint** (is a formula computed to help with sample size): If no and  $N < 50$ , allocate 0. If yes, allocate 2.

**Data collection:**

9) a) Was the method(s) of data collection described?  

No	Yes
1	2

b) Was the method appropriate given the:  
i. research question  

No	Yes
1	2

ii. nature of the data required (Will this data support the analysis?)  

No	Yes
1	2

**Data analysis:**

10) a) Was the analysis described?  

No	Yes
1	2

b) Was the analysis appropriate given the  
i) Research question  

No	Yes
1	2

**Results:**

11) Was statistical significance reported accurately?

No	Yes	Yes, with alpha levels
1	2	4

12) Is findings presented clearly and unambiguously?

No	Yes
1	2

13) Was a clear conclusion drawn?

No	Yes
1	2

14) Were appropriate recommendations made?

No	Yes
1	2

15) Did the authors identify and discuss limitations to the study?

No	Yes
1	2

**Ethics**

16) Was ethical approval obtained from an identifiable committee/ body?

No	Yes
1	2



UNIVERSITY of the  
WESTERN CAPE

**Section B: Instrumentation (Maximum score = 51).**

**Purpose of Instrument:**

1) Did they comment on the purpose of the measure?

No	Yes
1	2

2) Did they specify what attribute/ characteristic, construct will be measured?

No	Yes
1	2

3) Did they state whether the measure is to be used for screening purposes or in-depth diagnostic assessment?

No	Yes
1	2

- 4) Did they explain what type of decisions could be made on the basis of the test scores?
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 5) Did they specify for which population the measure is intended?
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |

**Methodology:**

- 6) Was the construct(s) theoretically defined? (did the researcher undertake a thorough literature study to define the construct)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 7) Was the construct operationally defined? (how construct will be measured, domains identified for measuring)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 8) Did they report on the methodology used to derive an operational definition (focus groups and individual interviews with various role players)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |

**Items & Administration:**

- 9) Did the researcher report on how they selected specific items?
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 10) Did the researcher comment on assembling of the items? (Arranging items, finalising length of test)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 11) Did the researcher comment on development of administration instructions?
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 12) Did the researcher pilot the test and items?
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 13) Were results evaluated in terms of item (difficulty, discriminating power, bias?)
- |    |     |
|----|-----|
| No | Yes |
| 1  | 2   |
- 14) Was there feedback on revision of test and item content?

- |  |    |     |  |
|--|----|-----|--|
|  | No | Yes |  |
|  | 1  | 2   |  |
- 15) Was there feedback on standardisation of administration procedures?
- |  |    |     |  |
|--|----|-----|--|
|  | No | Yes |  |
|  | 1  | 2   |  |
- 16) Did the reviewers comment on cultural, linguistic and gender appropriateness?
- |  |    |  |  |
|--|----|--|--|
|  | No | Yes  |  |
|  | 1  | 2 (one aspect), 3 (2 aspects), 4 (more than three aspects) |  |

**Validity:**

- 17) Were the items reviewed by means of experts for content validation?
- |  |    |     |  |
|--|----|-----|--|
|  | No | Yes |  |
|  | 1  | 2   |  |
- 18) Was the construct validity of the instrument tested statistically?
- |  |    |     |  |
|--|----|-----|--|
|  | No | Yes |  |
|  | 1  | 2   |  |
- 19) Were the psychometric properties of the final version established?
- i. *Validity:*
- |  |                      |                   |                      |
|--|----------------------|-------------------|----------------------|
|  | No                   | Yes               |                      |
|  | 1                    | 2 (face validity) | 3 construct validity |
|  | 4 criterion validity |                   |                      |

**Reliability:**

- 19) Were the psychometric properties of the final version established?
- i. *Reliability:* (internal consistency or test-retest or inter-rater reliability)
- |  |    |     |  |
|--|----|-----|--|
|  | No | Yes |  |
|  | 1  | 2   |  |

**Interpretation:**

- 20) Was a proper guide for interpretation developed?
- |  |    |     |  |
|--|----|-----|--|
|  | No | Yes |  |
|  | 1  | 2   |  |
- 21) How long ago was the test developed?
- |  |                |                |              |
|--|----------------|----------------|--------------|
|  | > 20 years ago | < 10 years ago | <5 years ago |
|  | 1              | 2              | 3            |
- 22) Is it clear that there might be more relevant assessment measures?
- |  |    |     |  |
|--|----|-----|--|
|  | No | Yes |  |
|  | 1  | 2   |  |







## Appendix E: Permission for SFS scoring system

Mr. G. Robertson

Department of Psychology

UWC

12 March 2015

### Re: Permission to use the SFS scoring system – Version D

Dear Mr Robertson

Thank you for your interest in using the SFS scoring system. I hereby give you permission on behalf of the collaborating authors to use the critical appraisal tool in your research towards the M.A.Psych degree. I would like to request that you provide us with feedback as to how you found the tool in your research. Your feedback will be valuable for future refinement

The SFS scoring system is currently being reviewed for publication. You can include a copy of the tool in your examination copy of the thesis provided that you insert a watermark on the appendix to indicate that it is not for reproduction. The final copy of your thesis that is uploaded into the library should not contain the critical appraisal tool. You can provide my contact details for anyone who is interested in using or reviewing the tools. This letter must be included as an appendix and the conditions stipulated reflected in your ethics section.

You can use the following references to support your thesis write up:

1. Smith, M.R., Franciscus, G. Swartbooi, C. Munnik, E. & Jacobs W. (2015). The SFS scoring system. In Smith, M.R. (Chair). *Symposium on Methodological Rigour And Coherence: Deconstructing The Quality Appraisal Tool In Systematic Review Methodology* conducted at the 21<sup>st</sup> National Conference of the Psychological Association of South Africa, South Africa.
2. Smith, M.R. (2015). Methodological Rigour and Coherence: A concept paper. In Smith, M.R. (Chair). *Symposium on Methodological Rigour And Coherence: Deconstructing The Quality Appraisal Tool In Systematic Review Methodology* conducted at the 21<sup>st</sup> National Conference of the Psychological Association of South Africa, South Africa.

The following references represent a sample of studies in which the scoring system and specifically version D was piloted

3. Trimble, L. & Smith, M.R. (2015) Strategies aimed at developing capacity in research supervisors. In Smith, M.R.. (Chair). *Symposium on Research Capacity Building: Identifying Elements From Supervision And Staff Development* conducted at the 21st National Conference of the Psychological Association of South Africa, South Africa.

4. Hendricks, A. Simons, A. & Smith, M.R. (2015). Strategies to develop research capacity in graduate students. In Smith, M.R.. (Chair). *Symposium on Research Capacity Building: Identifying Elements From Supervision And Staff Development* conducted at the 21st National Conference of the Psychological Association of South Africa, South Africa.
5. Simons, A. & Smith, M.R. (2015). Strategies to enhance research capacity in early career academics: A Systematic review. In Smith, M.R. (Chair). *Symposium on Research Capacity Building: Identifying Elements From Supervision And Staff Development* conducted at the 21st National Conference of the Psychological Association of South Africa, South Africa,.
6. Rae, N. & Smith, M.R. (2015).Demographic and personal factors that impact completion of student research. In Smith, M.R. (Chair). *Symposium on Research Capacity Building: Identifying Elements From Supervision And Staff Development* conducted at the 21st National Conference of the Psychological Association of South Africa, South Africa.

You can also cite the references of the unpublished theses of

Abigail Simons

Nicolette Rae

Lyle Trimble

Erica Munnik



I wish you well on your research and academic endeavours.

Sincerely



.....  
Dr. Mario R. Smith