VALIDATION OF AUTISM SCREENING ASSESSMENTS: COMPARISON OF THE SOCIAL COMMUNICATION QUESTIONNAIRE, SOCIAL RESPONSIVENESS SCALE AND 23Q WITH DSM-5 IN ASSESSING FOR AUTISM SPECTRUM DISORDER (ASD) IN UGANDA

By

Jorem Emmillian Awadu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Rehabilitation Counselor Education—Doctor of Philosophy

2021

ABSTRACT

VALIDATION OF AUTISM SCREENING ASSESSMENTS: COMPARISON OF THE SOCIAL COMMUNICATION QUESTIONNAIRE, SOCIAL RESPONSIVENESS SCALE AND 23Q WITH DSM-5 IN ASSESSING FOR AUTISM SPECTRUM DISORDER (ASD) IN UGANDA

By

Jorem Emmillian Awadu

The lack of validated autism spectrum disorder (ASD) screening instruments hampers rehabilitation efforts for children with ASD in Africa. In the present study, the psychometric properties of two ASD screening instruments, that have been well-validated in the United States (US) context, were examined in the Ugandan context. These instruments were the Social Responsiveness Scale-Second Edition (SRS-2; Constantino & Gruber, 2012) and the Social Communication Questionnaire (SCQ; Rutter, Bailey & Lord, 2003). These two instruments were cross-validated with the 23Questions Screener (23Q; Kakooza-Mwesige et al., 2014) - a developmental disability screening measure for assessing ASD symptoms among children in Uganda. Two distinct groups of study participants were recruited: (a) 51 caregivers of children aged 4 - 17 years with a known diagnosis of ASD and (b) 56 caregivers of typically developing children (non-ASD) of the same age range. Caregivers of children with ASD were recruited either during routine clinical follow-up visits at a national hospital or through a special needs school—with both sites drawing cases from the city of Kampala and the surrounding districts. Caregivers of typically developing children (i.e., children not identified or suspected of having a developmental disability) were recruited either during medical wellness visits or through schools within the same region. Results showed that the SRS-2 and SCQ were both reasonably reliable and valid for use in distinguishing Ugandan children with ASD from those who are typically developing. Summation of items 19 - 23 of the 23Q (typically used to screen children below five

years of age) yielded higher internal consistency reliability than a combination of all ten available 23Q ASD items (i.e., items 14 – 23). For discriminant validity, children with a diagnosis of ASD scored significantly higher than their non-ASD peers on all study instruments. Based on the sample results, the optimum SRS-2 cut-score was lowered to 62, while that of the SCQ was retained at 15. A cut-score of one or more on the 23Q, despite having high sensitivity yielded unacceptably low specificity. Increasing the 23Q cut-score to 3 or more improved specificity but lowered sensitivity albeit to a moderate range.

To my parents, Mrs. Jane Margaret Iyobu Akope (R.I.P.) and dad, Mr. John Ejou Akope.
For you, my wife, Esther Akwii, and son, Jethro Victor Awadu
To you, Dr. Kevin Aanyu, and all my siblings

ACKNOWLEDGEMENTS

First, I would love to thank Dr. Gloria Lee – my academic advisor. You constantly challenged me to be better. The world needs more professors like you.

Dr. Martin Volker, I do not know how best to say THANK YOU in a befitting way. Ours was a chance meeting on academic pursuits. However, you took me up wholeheartedly. Your guidance enabled me to succeed in this endeavor. You have created one of the few psychometricians in Africa.

To Drs. Michael Leahy and Andrew Nay, yes, we finally made this happen. It was an honor having you on my dissertation committee. Your questions, whether in the lift in Erickson Hall or during our meetings made this study possible.

Uniquely, I would like to thank Drs. Connie Sung, John Kosciulek, Hung Jen Kuo, Todd Lewicki, Su Pi, and Cary Roseth. Thank you for all the support over the years.

To my mentors: Drs. Catherine Abbo, Angelina Kakooza-Mwesige, Amara Ezeamama, and Michael Boivin, I am grateful for your guidance and support.

To Rebecca Akello, Annet Birabwa, Julius Caesar Ojuka, Wankya Ronald Kiduma, Robert Katende, and the administrators in Butabika National Mental Health Referral hospital and Special Children's Trust Uganda, thank you for making this study possible.

Finally, I would love to thank the administrators of the College of Education, Michigan State University. Your several research scholarships to me made this ambitious student research project possible.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Autism Spectrum Disorder

The fifth edition of the Diagnostic and Statistical Manual (DSM-5) defines Autism Spectrum Disorder (ASD) as a condition characterized by impairment in two domains: social-communication and interaction; and restrictive, repetitive behaviors, activities, or interactions (American Psychiatric Association [APA], 2013). In the Diagnostic and Statistical Manual, Fourth Edition, Text Revision (DSM-IV-TR), autistic disorder was defined by impairment in the three areas of social interaction, communication, and restrictive, repetitive, and stereotyped patterns of behavior (APA, 2000). Autism did not have its separate category under Pervasive Developmental Disorders, which included Autistic Disorder, Asperger's Disorder, Rett's Disorder, Childhood Disintegrative Disorder, and Pervasive Developmental Disorder Not Otherwise Specified (APA, 2000). In addition to the major recategorization of the five disorders, the diagnostic criteria of autism merged the social and communication subdomains, which were independent domains in the DSM-IV-TR. The repetitive restricted behaviors, activities or interactions, and the requirement that onset of impairments occur before the age of three years of age have held constant across the various iterations to the autism diagnostic criteria (APA, 2013; Lecavalier & Norris, 2010; Volkmar, 2012; Young & Rodi, 2013). In recent years, most researchers and practitioners have used the terms autism and autism spectrum disorder interchangeably to refer to the same disorder. Such interchangeable usage will occur in the context of the present study as well.

As of 2015, it was estimated that one in every 145 children worldwide had autism spectrum disorder (Presmanes Hill, Zuckerman, & Fombonne, 2015). Some researchers posit a prevalence rate of one percent and 1.5% for ASD in developing and developed countries

respectively. One in 54 children in the United States, by age the age of eight, has an ASD diagnosis, according to the most recent estimate from the Autism and Developmental Disabilities Monitoring [ADDM] Network (Maenner, Shaw, Baio, et al., 2020). Well-developed population-based prevalence studies for ASD in Africa are lacking, but researchers and practitioners suspect that prevalence may be higher than in developed countries.

For instance, Kakooza-Mwesige and colleagues (2014) estimated that 1.2 –1.3/100 children have ASD in Uganda. Although concerns about the sampling procedure put this estimate in doubt (Abubakar, Ssewanyana, & Newton, 2016), it nevertheless provides a preliminary, best available estimate of the prevalence in Africa. It is important to note that the estimates from Presmanes Hill et al. and Kakooza-Mwesige et al. occurred close in time to each other because evidence from developed countries suggests that the prevalence of ASD diagnosis has been increasing over time and that the most recent estimate in the U.S. from the Centers for Disease Control and Prevention (CDC) puts the U.S prevalence rate at 1 in 54 (Maenner et al., 2020).

Assessing for ASD and its Challenges in Africa

Researchers and practitioners agree that the lack of validated ASD screening assessments hampers the diagnosis and rehabilitation efforts of children with ASD (Abubakar et al., 2016; Franz, Chambers, von Isenburg, & de Vries, 2017). Most clinicians in Africa rely on the DSM-IV-TR (American Psychiatric Association, 2000) for ASD diagnosis, which although readily available, needs extensive clinical training and advanced education for one to administer appropriately.

In more developed countries, a clinical psychologist needs doctoral training to administer advanced diagnostic instruments such as the Autism Diagnostic Observation Schedule (ADOS)

(Lord et al., 2000), and the Autism Diagnostic Interview-revised (ADI-R) (Rutter, Le Couteur, & Lord, 2003), which are considered the gold standard for ASD diagnosis. Moreover, unlike most developing countries, diagnosticians in the developed world are utilizing the DSM-5 despite arguments for or against its use. Such advanced training is limited in Africa, is held by a few psychologists, or neurologists, and yet the number of people in need of their services is staggeringly high (Mullan et al., 2011). Besides, the widespread practice is that primary healthcare workers with limited training run rural health centers in Africa (Mullan et al., 2011). The few better-trained and equipped clinicians often live and work in the urban areas (Tekola et al. 2016).

The characteristics of ASD are closely related to other neurodevelopmental disorders. Without a precise differential screening tool, ASD is likely to be misdiagnosed for other disorders such as Attention Deficit and Hyperactivity Disorder (ADHD), Intellectual Disability (ID) or other closely related neurodevelopmental ASD comorbidities. The thin line between ASD and other neurodevelopmental disorders calls for good differential diagnostic skills, which only well-trained clinicians possess, as well as instruments. Subsequently, the risk of misdiagnosis may be high in Africa because of the limited number of practitioners trained to competently carry out a precise diagnosis of ASD. Given this problem, it is not surprising that ASD cases, especially in rural areas, are identified later during the child's development in Africa (Smith, Malcolm-Smith, & de Vries, 2017).

Furthermore, there is often a long waitlist before a confirmation of a diagnosis takes place for a child with ASD, let alone receiving services. A useful screening tool that does not require advanced level professionals will streamline the assessment process and better enable children that are at risk for ASD to receive appropriate services early. This practice is likely to help focus

the utilization of advanced professionals' efforts on conducting confirmatory diagnoses of ASD.

The need for early detection using a screening tool is critical as it will help with early identification of children at risk of ASD by practitioners with less advanced diagnosis training serving primarily in the rural areas, hence increasing the number of referrals for better diagnostic assessment. Referrals ensure that the highly specialized efforts wielded by the few well-trained professionals are focused on those identified as being at higher risk so that they can be referred to rehabilitation services promptly (Eaves, Wingert, & Ho, 2006).

Uganda is among the few African countries where there has been an attempt at validating a culturally sensitive screening instrument, which caters for children at risk of ASD and other neurodevelopmental disorders like epilepsy, cerebral palsy, and cognitive impairment, among others (Kakooza-Mwesige et al., 2014). There was a limited attempt at validating the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) (Smith, Malcolm-Smith, & de Vries, 2017) in South Africa with a small sample size ($N = 7$). It was translated into the Afrikaans version to generate guidelines to enhance cultural sensitivity and appropriateness. Kakooza-Mwesige and others (2014) stressed the need for further work in the validation of autism-related screening instruments in low resource countries to assist in the identification of affected individuals and boost rehabilitation efforts.

Purpose of the Study

Without deliberate efforts to identify ASD assessment and screening instruments that fit the local context, we run the risk of missing the benefits accruing from early diagnosis and management of ASD in African countries, including Uganda. It is likely that a substantial proportion of children with ASD will not receive rehabilitation services. This poses long-term negative impact in their adulthood, as well as overall economic impact because of additional

services needed for impacts that could have been alleviated or reduced. The primary purpose of this study, therefore, is to investigate the psychometric properties of two commonly used, well-validated ASD screening instruments used in the US with children in Uganda. The Social Responsiveness Scale, Second Edition (SRS-2; Constantino & Gruber, 2012) and Social Communication Questionnaire (SCQ; Rutter, Bailey & Lord, 2003) will be assessed for utility in the screening of ASD symptoms for children aged 4 -18 years in Uganda. The ten ASD specific items of the 23Questions screener (23Q; Kakooza-Mwesige et al., 2014), an instrument developed in the African context to screen for a broad array of neurodevelopmental disorders, including ASD, will be used to examine concurrent and discriminant validity.

Potential Implications of the Study

The results of this study will provide an initial step in understanding the utility of the SCQ and SRS-2 in applying to children at risk of ASD in the Ugandan culture. The study will have significant application to the identification of valid, reliable, and efficient ASD screening tools, which will aid in the early identification of those at risk of ASD. The broader community of professionals, such as teachers, counselors, social workers who may not have advanced ASD identification training such as in medicine and psychology will also use them. Combined with ease of use by front-line clinicians, it will allow for improved early screening and identification of at-risk children and help free up time for pediatricians, and psychologists to conduct formal ASD diagnoses. Children diagnosed with ASD will then be more likely to be referred for appropriate services with delay—improving the likelihood of better adult outcomes. Therefore, results of this study could potentially lead to the validation of the first tool specifically adapted for ASD screening in Uganda. This will improve the likelihood of early identification and access to appropriate services for children with ASD and their families.

Research Questions

1) In the Ugandan context, do total scores for the three ASD instruments (i.e., SCQ, SRS-2, and 23Q) yield adequate internal consistency relative to values reported in the test manual or available studies for those with ASD?
2) In the Ugandan context, do total scores for the three ASD instruments (i.e., SCQ, SRS-2, and 23Q) yield adequate internal consistency relative to values reported in the test manual or available studies for non-ASD cases?
3) In the Ugandan context, do total scores from the three instruments (i.e., SCQ, SRS-2, and 23Q) yield substantial mean differences consistent with construct-related (i.e., ASD construct) differentiation between ASD and non-ASD groups?
4) In the Ugandan context, will the use of the recommended screening cut-scores for each of the three instruments (i.e., SCQ, SRS-2, and 23Q) result in adequate sensitivity and specificity using DSM-5 ASD diagnosis vs. non-ASD cases as the outcome variable?
5) Beyond the recommended cut-scores for screening reported in each test manual, are there more optimal cut-scores for each of the three instruments (i.e., SCQ, SRS-2, and 23Q) in the Ugandan context?
6) Are there significant convergent relationships among the total scores of the three instruments (i.e., SCQ, SRS-2, and 23Q) in the Ugandan context?

CHAPTER 2: LITERATURE REVIEW

Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a disorder that is marked by impairment in social communication or social interaction as well as restricted, repetitive patterns of behavior, interests, and activities in individuals (APA, 2013). Leo Kanner first reported autism in 1943, where he referred to it as infantile autism. According to Kanner, children who had infantile autism tended to isolate themselves from others, had limited language abilities and enjoyed doing a specific activity repetitively. This profile mimics the description of a child with autistic disorder, according to the fourth edition of the Diagnostic and Statistical Manual, Text Revision (DSM-IV-TR). In 1946, Hans Asperger reported another group of boys with similar characteristics who appeared to be higher functioning. Cases who demonstrated this pattern, which Asperger referred to as having autistic psychopathy of childhood, later came to be referred to as Asperger's Disorder (Miller & Ozonoff, 1997).

Kanner is credited with being the first to call the condition autism. Both Kanner and Asperger described children with unique behaviors although according to Asperger, Kanner's autism was more severe as compared to Asperger's syndrome, which was hard to diagnose especially in childhood (Wing, 1986). Kanner's syndrome came to be referred to as autism while that of Asperger was referred to as Asperger's syndrome. Not much was documented about autism in the English-speaking world because the original descriptions of children with autism were in German. It was not until Van Krevelen (as cited in Wing, 1986) whose paper was urging for the separation of Asperger's syndrome from Kanner's autism/syndrome was published. Lorna Wing specifically introduced Asperger's syndrome to the English-speaking scientific community

by describing a series of cases like those of Asperger (Klin & Volkmar, 2005; Woodbury et al., 2005).

There have been heated debates about the clinical difference between autism and Asperger's syndrome as provided for in the DSM-IV. Most clinicians argued that making a clinical diagnosis of Asperger's syndrome under the DSM-IV was difficult (Woodbury et al., 2005). The significant difference between autism and Asperger's syndrome was associated with the intact language skills of those with Asperger's syndrome as well as the late age of its onset. It was, therefore, challenging to make a conclusive diagnosis in situations where a parent or other primary caregiver was not accurate on developmental milestones of the child (Klin & Volkmar, 2005). Researchers re-evaluated the four cases described by Asperger as having Asperger's Syndrome using the DSM-IV and showed that they would be re-assigned to a diagnosis of autism and not Asperger's Syndrome as initially indicated (Miller & Ozonoff, 1997). Such sharp research and clinical disagreements have been the major contributors to the merging of the different autism subdomains in one, i.e., autism spectrum disorder in the current DSM-5.

Etiology and Prognosis

Autism Spectrum Disorder etiology has been associated with both genetic and environmental factors, but not a single biomarker has been identified for the diagnosis of one with ASD (Bello-Mojeed, Omigbodun, Bakare, & Adewuya, 2017). Parental history of mental illness, mother's medication with psychotropic drugs in the early stages (first trimester) of pregnancy, bearing children while of advanced age (males 50 years and older, and females 40 years and older), genetic factors, environmental toxins are considered potential risk factors for ASD (APA, 2013). However, genetic factors have the most robust empirical support for ASD

etiology thus far. Research in twin concordance rates for autism indicates hereditary estimates between 37% to above 90% (Geschwind, 2011).

The age of onset for ASD is between 12 to 24 months (about 2 years). Some children are, however, diagnosed later beyond three years of age (APA, 2013; Meng et al., 2013). Research that enables the identification of younger children at risk of ASD has been ongoing. Children as young as six months old can now be screened for ASD, although most accurately diagnosed at two years when apparent deficits can be ascertained. ASD onset is characterized by either rapid or gradual child developmental regression in language, social interaction and restricted, repetitive patterns, behavior, or activities (APA, 2013).

More boys than girls are diagnosed with ASD, with a male to female ratio of between 4:1 and 5:1, and most females diagnosed with ASD in clinical samples having co-occurring intellectual disability (APA, 2013). Symptom masking, support, and compensation, especially as related to societal gender roles, may mask the deficits in social abilities among higher-functioning females with ASD (APA, 2013; Lai, 2015). Thus, detection and diagnosis are more difficult in females (especially those who are high functioning).

Autism is a lifelong disorder although those diagnosed with it currently have a better outlook to life when compared to more than 50 years back (APA, 2013; Lord et al., 2018; Meng et al., 2013). More individuals with ASD can work and live independently in the community depending on the level of severity. Those whose language and intellectual abilities are less affected (non-impaired or intact) can function better in adulthood. Early identification and interventions have been associated with better adult outcomes (Lord et al., 2018).

Global Burden

Autism prevalence rates are projected to be increasing worldwide (APA, 2013; Sharma, Gonda & Tarazi., 2018). Greater awareness, better diagnosis, and screening instruments, and change in the diagnostic criteria from the DSM-IV-TR to the DSM-5 are some of the reasons advanced for the increasing prevalence (Magaña & Vanegas, 2017).

Having a child with ASD has been associated with poor mental and physical health outcomes for parents. Parents of children with ASD often report higher levels of stress and depression resulting from constantly worrying about their child (Lee et al., 2009). Diagnosis, rehabilitation, and service needs of children with ASD are expensive and often lead to many negative outcomes, such as spousal conflicts (Gona et al., 2015; Schlebusch, Dada, & Samuels., 2017). This results in poor quality of life for parents. This is further worsened by negative societal attitudes and the limited number of professionals capable of providing the services needed for the functioning of children with ASD. Parents have a few resources for respite care and as a result isolate themselves, which further increases the risk for stress. The much needed social and professional support sought by parents could be affected, hence, negatively affecting the rehabilitation efforts for the child with ASD (Fewster & Gurayah, 2015; Franz, Chambers, von Isenburg, & de Vries, 2017; Gona et al., 2015; Kakkar & Srivastava, n.d.).

Autism Diagnosis

The first clinical documentation of the criteria to diagnose autism was captured in the Diagnostic and Statistical Manual of Mental Disorders, Third Edition [DSM-III] in 1980 and the International Classification of Diseases [ICD-10]. In the DSM-III, autism was recognized as a unique disorder in children with no known etiology, but with a different clinical presentation. There have been several changes in the categorization of autism right from the first time it was

classified for clinical purposes in the DSM-III-TR (DSM-III, APA, 1980). Research in autism increased as a result and led to changes in the descriptions and diagnostic criteria in the subsequent versions of the DSM. For instance, in the DSM-IV-TR, autism was categorized under Pervasive Developmental Disorders (PDDs) with other related diagnostic entities. The five PDDs were Autistic Disorder, Asperger's Disorder, Rett's Disorder, Childhood Disintegrative Disorder (CDD) and Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS), a designation for cases that did not meet full criteria for another PDD but shared significant core symptoms (APA, 2000). In the latest revision of the Diagnostic and Statistical Manual for Mental Disorders, Fifth Edition (DSM-5), the prior diagnostic designations of autistic disorder, Asperger's disorder, and PDD-NOS under DSM-IV-TR were re-organized into the single diagnostic category of Autism Spectrum Disorder (Magana & Vanegas, 2017).

The merging of three of the PDDs into one diagnosis of ASD in the current DSM-5 has been met with mixed reactions. Some self-advocates, parents of children with ASD and the concerned community at large, argue that the DSM-5 changes reduce the chance that a child is diagnosed with autism, hence impacting the child's ability to access disability services through disability insurance (Magana & Vanegas, 2017). Others argue that the change will increase the likelihood of reliable and accurate diagnosis (i.e., reduce the likelihood of missed or false negative cases). For example, Magana and Vanegas (2017) showed that the DSM-5 increased the likelihood of children of Hispanic descent being diagnosed with ASD when compared to the DSM-IV-TR. These authors indicated that the DSM-5 identified more children with ASD of Spanish speaking parents, with the social communication and social interaction domains being significantly different for the ASD and developmental disability groups under the DSM-5

revision. This was not the case under the DSM-IV for the same demographic group (Magaña & Vanegas, 2017; Vanegas, Magaña, Morales, & McNamara, 2016).

Despite the different views on the DSM-5, autism diagnosis takes advanced training for one to appropriately utilize the various diagnostic and screening tools. Often, these advanced diagnostic assessments rely on a combination of parental reports on child developmental history or milestones, and child observation by multidisciplinary teams (pediatric psychiatrists, psychologists, counselors among other professionals) (Lord et al., 2018).

Advantages of Autism Screening Tools

For children living in developing countries like Uganda, an ASD diagnosis is often tedious and expensive with little guarantee of an accurate diagnosis, especially for those living in rural areas (Bakare & Munir, 2011; Bello-Mojeed et al., 2017; Mazurek, Curran, Burnette, & Sohl, 2019). Professionals trained in advanced autism assessment are scarce, with most living in urban areas, hence reducing the chances for early diagnosis for children and family living in the rural settings due to inaccessibility issues. Autism screening, however, does not require advanced training for one to administer. This reduces diagnosis delay for individuals from a low socioeconomic status as well as costs to parents, as screening tools are less expensive than advanced diagnostic assessment. Also, timely and appropriate referrals lessen the effect of potential societal stigma to both parents and their children through timely access to services and support by more knowledgeable professionals (Kakooza-Mwesige et al., 2014; Mazurek et al., 2019; Woolfenden, Sarkozy, Ridley, & Williams, 2012).

Self-rating scales such as the Social Communication Questionnaire [SCQ] (Rutter, Bailey & Lord, 2003), Social Responsive Scale, second edition [SRS-2] (Constantino & Gruber, 2012) and others that are administered to parents or primary caregivers of children often reduce the

time required of the clinicians to make a diagnosis. They are cost-effective, less complicated to administer and can identify those in need of further clinical assessments.

Challenges Faced in the Utilization of Autism Screening Tools

Finding a uniform autism screening tool catering for all cultural and demographic variance is complicated. Most are developed in North America and not validated for use in African cultures (Durkin et al., 2015). Such unvalidated instruments are nonetheless utilized in ASD screening by professionals in some African countries. The possibility of misdiagnosis stemming from these unvalidated tools is therefore high. False positives resulting from using non-validated screening tools could lead to inappropriate intervention referral, stigma, anxiety, and extra financial burden on the parents as well as the healthcare system (Marlow et al., 2019). Further, younger children with less visible ASD characteristics are hard to identify especially where primary healthcare providers are not well trained. Autism screening instruments are equally expensive, and demand a better trained human resource, which is scarce in Africa. Professionals with limited ASD training often fill the human resource gap (Marlow, Servili, & Tomlinson, 2019).

Autism Diagnostic and Screening Instruments

Diagnosis of ASD involves a combination of procedures. In psychology, behavioral observation, and taking of developmental history from primary caregivers of the person being assessed form the most reliable basis for an autism diagnosis. Below are highlighted standard diagnostic and screening instruments used in the identification of autism.

Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)

According to the DSM-IV-TR, to be diagnosed with autism, one must meet the set criteria for autistic disorder. Specific features outlined in the three core areas, i.e., A, B and C as delineated in the DSM-IV-TR must be met. One must have a total of six or more items across all the sub-categories (A, B, or C) with at least two from sub-category: A) qualitative impairment in social interaction, and one under both B) qualitative impairments in communication and C) restricted repetitive and stereotyped patterns of behavior, interests, and activities (APA, 2000).

For qualitative impairment in social interaction (criteria A), one must have at least two or more features of the following: 1) marked impairment in the use of nonverbal behaviors like eye-to-eye gaze, facial expression, body postures and gestures to regulate social interaction. 2) failure to develop peer relationships that are appropriate to one's developmental level. 3) lack of spontaneous seeking to share interests or enjoyment, and 4) a lack of social or emotional reciprocity (APA, 2000).

For qualitative impairment in communication (criteria B), one must have at least one of the following: 1) a delay or a total lack of the development of spoken language (not accompanied by alternative modes of communication like gestures. 2) a marked impairment in the ability to initiate and sustain a conversation with others (among those with adequate speech). 3) repetitive or stereotyped use of language (echolalia), and 4) lack of varied, spontaneous make-believe play that is appropriate to their developmental level (APA, 2000).

A person must meet at least one criterion for the Restricted Repetitive and Stereotyped Behavior (criteria C): 1) intense or abnormal preoccupation with one or more interest(s). 2) inflexible adherence to routine or rituals. 3) stereotyped and repetitive motor mannerisms (hand

or finger flapping or twisting, or complex whole-body movements, and 4) persistent preoccupation with parts of objects.

It is important to note that to be diagnosed with autistic disorder, evidence of delay or abnormalities manifested before three years of age is required. Also, the symptom presentation should not have been better explained by a diagnosis of Rett's Disorder or Childhood Disintegrative Disorder (APA, 2000).

The Asperger's Disorder diagnosis is given to individuals who do not have a delay in their language development, but with sustained impairment in social interaction, development of restricted, repetitive patterns of behavior, interests, and activities. These must cause significant impairment in social, occupational, or other areas of daily functioning. While for a diagnosis of PDD-NOS, one must have pervasive impairment in social interaction and either the presence of stereotyped behavior, interests and activities, or substantial communication impairment(s), or both but not meet the full criteria for autistic disorder or Asperger's disorder. This might occur due to a late age of onset, some symptoms being subclinical or insufficient to meet criteria for another PDD, unusual variation in symptomology, etc. (APA, 2000).

Diagnostic and Statistical Manual for Mental Disorders, Fifth Edition (DSM-5) Criteria

According to the DSM-5 (APA, 2013), to be diagnosed with an autism spectrum disorder, one must meet the diagnostic criteria established by having deficits in all the two outlined criteria, that is: 1) social communication and social interaction domain (criteria A), and 2) at least two of the four outlined criteria under the restricted, repetitive sub-domain (criteria B).

Criterion A delineates the social communication and interaction aspects of a person with autism, which includes: 1) deficits in social-emotional reciprocity, 2) difficulty in nonverbal

communication behaviors, and 3) deficits in developing, maintaining and understanding relationships (APA, 2013).

Criteria B pertains to the restricted and repetitive behavior, interaction and activities subdomain of a person suspected to have autism. Under this subdomain, one must have any two of the four criteria outlined: 1) stereotyped or repetitive motor movements or use of objects, 2) insisting on sameness, inflexibility in routines, or ritualized patterns of verbal or nonverbal behavior, 3) highly restricted, fixated interests that are abnormal in intensity or focus, and 4) hyperactivity to sensory input or unusual interest in sensory aspects of the environment (APA, 2013).

The deficits outlined above must have been present in the early developmental period. However, such difficulties may not be fully manifest until social demands exceed the limited capacities of the individuals and may be masked by learned strategies later in life (APA, 2013). Also, the symptoms must cause severe impairment in social, occupational, or other areas of the person's current functioning. The outlined disturbances ought not to be better explained by intellectual disability (intellectual developmental disorder) or global development delay (APA, 2013).

Specifiers. The DSM-5 goes further in delineating different specifiers of symptom severity utilized in the diagnosis of individuals with autism. The various levels in the specifiers include:

Level 1: In this level, such a person will have difficulty in the social communication domain. Individuals with this specifier require support. Without the necessary support in place for them, the difficulty in this domain makes their impairments noticeable. They experience difficulties in initiating and sustaining social interaction. For example, such individuals may be

able to speak in full sentences, engage in communication with others, but their back- and forth interaction with others fails. Their attempts to make friends are unsuccessful and odd. For the restrictive, repetitive behaviors domain, the individual is inflexible in their behavior, and it significantly interferes with their functioning in one or more contexts. This behavioral rigidity makes it hard for them to switch between activities. They also experience difficulties with planning, which hinder their independence (APA, 2013).

Level 2*:* This level is characterized by increased difficulty in both verbal and nonverbal communication skills. Unlike in level 1, even with the availability of support, social impairments are evident in the person. Self-initiated social interactions are limited in nature, and they have odd or reduced responses to social interaction from others. Such a person will, for example, speak simple sentences and their interaction with others limited to a specific narrow particular interest and a markedly odd nonverbal communication. The rigidity in behavior or other restricted, repetitive interests appears often enough in various settings to be evident to a casual observer. They experience distress and difficulty changing focus or action. These individuals require substantial support to function (APA, 2013).

Level 3*:* For one to receive level 3 for a specifier, he or she must require very substantial support. People with this specifier have immense difficulties in verbal and nonverbal communication, which impairs their interaction with others. Their initiation, ability to sustain or respond to the interactions or conversation with others is very limited. They have less intelligible speech, few words and rarely initiate interaction. Even when they do, it is to meet their own needs (APA, 2013). They mostly respond to straightforward social approaches by others. Their behaviors are not flexible, and they have incredibly high difficulty coping with change and other

restricted or repetitive behaviors. This remarkably affects their functioning in all contexts. They also experience great difficulty in changing their focus or action.

Severity specifiers (i.e., levels 1, 2 and 3) are used to describe the current symptoms of the person being diagnosed. However, in assigning individuals specifiers, professionals should be cognizant of the fact that ASD severity may vary depending on where (context) and when (time) in which the behavior unfolds.

Autism Diagnostic Interview, Revised

The Autism Diagnostic Interview-Revised (ADI-R) (Rutter, Le Couteur, & Lord, 1994) is an instrument administered to parents or other primary caregivers of a child suspected to have autism by professionals knowledgeable in autism symptoms and presentation. It has two forms, that is, the summary and lifetime forms. The interview entails having a parent or caregiver of a child answer questions about symptoms related to social communication, interaction and restrictive, repetitive activities and behaviors (Lord et al., 2000). Lord and colleagues published the ADI-R in 1994 after modifying the Autism Diagnostic Interview in 1989 (Lord et al., 1994). The ADI-R administered to parents of children with cognitive abilities corresponding to a developmental level of at least two years. The instrument has 93 yes or no items that gather background as well as current information in the three areas of language and communication, reciprocal and social interaction as well as restricted, repetitive, stereotyped behaviors in the child.

The ADI-R scores range from 0 (behavior not-present at all), 1 (behavior mildly present, 2 (behavior significantly present), to 3 (behavior severely present). The severity code of 3 is treated as 2 to reduce the undue weight that could be placed on individual items (Lord et al., 1994). During its administration, the interviewer can ask follow-up questions about the child to

appropriately score the behavior. The scale scores are totaled and coded by the test administrator. Administration of the interview takes between 1.5 to 2.5 hours to complete. It provides five different algorithms, i.e., two utilized for diagnostic purposes and three to assess current functioning as well as help with treatment planning. A child with recommended cut-scores in the three domains, that is, social reciprocity (10), communication (8 or 7 for verbal and non-verbal cases respectively) and restrictive repetitive behaviors (3) is considered to have autism. Also, these behaviors should have manifested in at least one of the subdomains before three years of age.

The ADI-R was initially validated on a sample of 20 children (i.e., $n = 10$ with an autism diagnosis and $n = 10$ without autism, but with a mental handicap or language delay). They ranged in age from 30 to 39 months. The ADI-R scores yielded intra-class correlation coefficients ranging from 0.93 to 0.97 and test-retest reliability estimates of 0.93 to 0.97 among interviewers. Further ADI-R validity evidence was gathered through participation of another group of parents ($N = 50$) who were seeking help for their preschool children. Among these parents, $n = 25$ had children with autism and a second group of $n = 25$ had children with cognitive delay or language impairment, but without autism. Their children's chronological age ranged from 36 to 59 months with mental age ranging from 21 to 74 months. Out of the total sample, interrater reliability was assessed using a subsample of 20 children (i.e., 10 with autism and 10 without autism) with multiple raters. This inter-rater analysis yielded kappa coefficients ranging from .63 to .89 for each item. Intra-class correlations were above .92 for all subdomain and domain scores.) Only one of the 25 children with autism failed to meet the algorithm criteria for autism under the ADI-R. Two of the 25 children without autism, but with mental handicap and language, yielded ADI-R scores consistent with autism. Thus, the ADI-R criteria for autism,

as outlined by Lord and colleagues (1994), provided good discrimination between the two groups (Rutter, Le Couteur, & Lord, 2003).

Application of the ADI-R has occurred in other linguistic and cultural contexts. For example, in a study that utilized a Spanish version of ADI-R, Magaña and Smith (2013) compared 48 Latino adolescents and adults who had ASD with a matched sample of non-Latino Whites using the summary and lifetime scores on the ADI-R. They found no significant difference between the two groups on total impairment of social reciprocity or communication; however, the Latino group had lower levels of restricted and repetitive behaviors compared to their White counterparts. Eight of the Latino and three of the White children did not meet the autism cut-score for the repetitive behavior domain with an interrater reliability of .88 for the interviewers. According to the study authors, this might have happened because Latino mothers often place more emphasis on interpersonal social development as compared to goals related to the development of autonomy among their children (Magaña & Smith, 2013).

Tsuchiya and others (2013) investigated the psychometric properties of a Japanese version of the ADI-R (i.e., the ADI-R-JV). To establish interrater reliability, they utilized a total sample of 51 individuals aged 2-19 years. These were divided into two groups: 1) 35 individuals who were clinically referred for ASD, and 2) 9 control group with no ASD diagnosis. For discriminant and diagnostic validity, the authors recruited 317 individuals aged 2-19 years and divided them into three groups: 1) n = 138 with autistic disorder (AD), 2) n = 28 PDD-NOS, and 3) n = 90 with non-ASD. Intra-class correlations of 0.80 were obtained for the three different domains for interrater reliability. For discriminant validity, the mean scores for those with AD were higher than the other two groups. The sensitivity and specificity of the ADI-R-JV were

reported at 0.92 and 0.89 respectively. However, the sensitivity of those below five years of age was 0.55 (Tsuchiya et al., 2013).

Another study with the Polish version of the ADI-R (ADI-R-PL) established psychometric validation of the assessment for use in Poland (Chojnicka & Pisula, 2019). The study consisted of 125 individuals aged two years old and above. They were divided into two groups, 1) n = 65 individuals diagnosed with ASD and 2) n = 60 controls, among whom 18 had non-spectrum disorders and 42 were typically developing children. The norming sample comprised a total of 178 individuals (n = 118 with ASD and n = 60 with no ASD). Intra-class correlations for children of two years old and above ranged from .96 -1.00 while the test-retest reliability for the various domains was .88- .91 for those with two years and above. Cronbach alpha for all domains ranged from .85 to .95 except for the areas of reciprocal peer interaction (.64) and restrictive repetitive behaviors (.63). The authors also reported the Comparative Fit Index (CFI) for the ADI-R-PL at .88 and .89 (slightly lower than the recommended estimate of .90 and above) for the two and three factor model, respectively. A root means square error of approximation (RMSEA) of .08 was established for both models, which is above the recommended .06 to indicate a good model fit. They concluded the three-factor model following the norming study by Lord (1994) for the ADI-R-PL.

In summary, evidence suggests that the ADI-R is a good diagnostic tool, but also one that requires substantial training to administer, score, and interpret appropriately. It takes a long time to administer and requires extensive training. The ADI-R should not be used as a single tool for diagnostic purposes, but with other complimentary assessments such as the ADOS-2 (Falkmer et al., 2013). Limited studies were published in the use of the ADI-R in other languages or cultures.

Initial results appeared to be mixed in terms of its psychometrics when applied to diverse cultures. Additional studies are warranted.

Autism Diagnostic Observation Schedule-2 (ADOS-2)

The Autism Diagnostic Observation Schedule-2 (ADOS-2; Lord et al., 2012) is a semi-structured assessment that entails observation of the child's behaviors or activities to confirm or disconfirm a potential diagnosis of ASD (Lord et al., 2012). Assessment using the ADOS-2 entails activities requiring the child or person being assessed to elicit behaviors characteristic of people with ASD. It is a standardized assessment of communication, social interaction, play or imaginative use of materials as well as restricted and repetitive behaviors for individuals suspected of having ASD.

Administration of the ADOS-2 entails observation of a person's behavioral presentation, coding, and scoring them using the scoring algorithm provided by the test providers. The choice of the module to be administered depends on one's expressive language skills, chronological age, and appropriateness of one's maturity or developmental level. The ADOS-2 has five modules: Module 1: children with no language skills; Module 2: children with minimal language skills (are not verbally fluent); Module 3: children and young adults with phrase speech; Module 4: older adolescents and adults and Module 5: a toddler module, which is utilized for toddlers aged 12 - 30 months (about 2 and a half years). Modules 1 and 2 can be administered to adolescents or adults with expressive language levels meeting the two modules although their interpretation must be made with caution because the instrument validation samples did not include adults or adolescents (Kamp-Becker et al., 2018). It takes between 40 to 60 minutes to administer each module with the possibility of scores ranging from 0 (the behavior targeted is not present) to a score of 2 or 3 (indicating the presence of a targeted behavior). The ADOS-2 was found to have

an inter-rater reliability ranging from .78 - .98 (for modules 1- 3) among research trained assessors (Lord et al., 2012) and a median inter-rater reliability of .74 - .84 (for all the four modules) among clinically trained ADOS users conducting clinical work across 13 sites in Sweden (Zander et al., 2016).

Limited studies have been conducted in the use of the ADOS in other cultural contexts. For instance, the ADOS-2 (ADOS-2-PL) was validated for use in Polish using a sample of 407 study participants. They were divided into two groups, 1) n = 193 with ASD, and 2) n = 208 without ASD (78 with non-autism spectrum and 130 typically developing). Its test-retest estimates ranged from .71- .95 for all algorithms except for the restrictive repetitive behaviors domain score which was lower for all the three algorithms in the "aged 5 years and older" algorithm of module 2, which was .41, module 3 (.54) and module 4 (.65). The internal consistency of all modules in the ADOS-2-PL was above .70 for all modules except module 3, which had 0.68. The sensitivity was over .90 for all modules, except for the "Aged 5 years or older" algorithm in module 2, which was 0.84. For specificity, all modules were above 0.80 except for in module 2 and 4 "aged 5 years or older" where it was .70. The mean scores of ASD were higher in comparison to those without ASD (Chojnicka & Pisula, 2019).

A 'pre-pilot' version of the Afrikaans-translated ADOS-2 was used to assess 47 children. Forty were from the community while seven were clinical samples referred to Red Cross War Memorial Children's Hospital Developmental clinic in South Africa (Smith, Malcolm-Smith & de Vries., 2017). Four of the seven children referred to the clinic had prior clinical diagnoses of ADHD, intellectual disability, global development delay and ectodermal dysplasia-cleft syndrome. The inclusion criteria of all participants were: being colored, ability to speak Afrikaans, from a low SES and having a child aged 1-20 years. All participants were divided into

two groups. Group 1 consisted of 24 caregivers with children under 10 years of age, with 20 parents from the community sample and four from the clinical sample. Group 2 consisted of 23 caregivers whose children were 10 or more years old, with 20 parents from the community sample and three from the clinical sample. The study authors reported limited psychometric properties for the instrument. However, they pre-piloted the translated ADOS-2 (Afrikaans version) with four caregivers from the clinical sample and the examiner.

The community samples were utilized for the validation of the appropriateness of social interactions and activities. For the clinical sample, language, activities, and cultural appropriateness of the ADOS-2 Afrikaans version were performed. Caregiver comments about the instruments were documented and thematically analyzed. Children and caregivers were not familiar with certain items, including the bathtub, remote controlled bunny, or baby dolls, among others. They were, however, familiar with plastic cups, plates, and utensils clearly indicating the inappropriateness of some test items. Such discrepancies affect the clinical validity as deductions of scores would have been made in such contexts. The ADOS-2 Afrikaans version by Smith and researchers (2017) provides an initial exploration of the application of the ADOS-2 in the African culture.

In summary, the ADOS is often used by well-trained professionals to make ASD diagnostic decisions (Lord et al., 1994) and is considered a “gold standard’ in diagnosing for autism. Its potential application in the African context, in translation and in terms of cultural appropriateness, is currently being explored.

Both the ADOS and ADI-R are reliable instruments in distinguishing children with and without ASD. However, for proficiency in its administration, one requires considerable training. Also, it is time-consuming to administer both instruments (Falkmer et al., 2013). Their prices are

equally prohibitive to most professionals and yet one must continuously purchase additional copies from the test authors. Autism screening tools are suitable alternatives because they are cheaper, require less training to administer, and are not time-consuming and yet with acceptable to high reliability as shown below.

Social Communication Questionnaire (SCQ)

The Social Communication Questionnaire (SCQ) was initially referred to as the Autism Screening Questionnaire (ASQ; Berument, Rutter, Lord, Pickles, & Bailey, 1999). The SCQ cannot be used to diagnose ASD in isolation but is useful as a screener for identifying those in need of further assessment by professionals who are appropriately trained and equipped to diagnose autism. It takes about ten minutes to fill out and five minutes to score. The SCQ has been validated for use with individuals whose chronological age is over four years and have a mental age of at least two years (Gau et al., 2011; Lord et al., 1994).

The items for the SCQ were adapted from the diagnostic scoring algorithm of the ADI-R, which incorporate content that is also consistent with the DSM-IV diagnostic criteria for autistic disorder (APA, 1994). There are two different forms of the SCQ, that is, the Lifetime and Current forms. Both consist of 40 items although item wording is slightly different across the forms to reflect the purpose and timing of the assessment. The Lifetime form is intended for diagnostic screening and the Current form is intended for assessing change over time by reflecting behavior from the most recent three-month period (Rutter, Bailey, & Lord, 2003). Only the Lifetime form was used in the present study.

The SCQ Lifetime form, referred to simply as the SCQ going forward, is a 40-item instrument filled out by parents or caregivers most familiar with the child suspected of having ASD. All items are responded to using a dichotomous (i.e., yes/no) scale (Berument et al., 1999;

Gau et al., 2011). Item 1, regarding the presence or absence of phrase speech", is not included in the total score calculation. However, one's response to item 1 determines which other SCQ items will be included in the calculation of the total score. For those with sufficient language, their SCQ total score can range from 0 to 39, based on the summation of items 2 through 40. For those without sufficient language, the total score can range from 0 to 33, as only items 8 through 40 are included in the calculation of the total. An optimal cut-score of 15 or more is utilized to differentiate ASD from other diagnoses. A slightly lower cut off is suggested, but not specified, for those with a substantial language impairment or other risk factors. Those under this condition are referred for further evaluation by more experienced professionals for a final diagnostic determination (Rutter et al., 2003).

Only the SCQ total score has been validated for screening purposes. However, the measure contains subsets of items reflecting each of the three domains of the ADI-R and the three core symptom domains of autistic disorder under DSM-IV. Though scores can be derived for these domains, information about domains is considered more important for purposes of SCQ's content validity and not for use in domain-specific screening. The first domain reflects reciprocal social interaction and consists of 15 items asking about, for example, the child's social smiling, interest in other children and offering comfort to others. The second is about communication issues and contains 13 items. This domain includes items asking if the child uses conventional gestures, reciprocates while in conversation and whether he or she has stereotyped utterances, etc. The third domain assesses restricted, repetitive, and stereotyped patterns of behavior. It consists of eight items which ask about such things as stereotyped body, hand, or finger mannerisms; ritualistic behavior; circumscribed interests and unusual pre-occupations by

the child; etc. Also, the SCQ has an item related to self-injurious behavior, as well as another on current language functioning (Berument et al., 1999; Rutter et al., 2003).

According to Rutter and researchers (2003), the primary standardization of the SCQ was with a total of 200 individuals who were participants of previous autism-related studies. The studies entailed participants with the various subdomains of autism under the DSM-IV together with those who had non-ASD disabilities often utilized for comparison purposes. The total number of cases utilized for analysis of the SCQ total score purposes was 214 (i.e., 177 males and 37 females) with ages ranging through two to 18 years. According to the parents, a total of 157 children had the verbal ability while 57 did not. A total of 213 children had their prior diagnostic information available. The 157 children who had verbal or language ability were divided into three groups. The first group were those with a diagnosis of autism (n = 71). The second group were those with ASD (n = 49) including Asperger's syndrome, PDD-NOS although it did not have those with Rett's syndrome nor childhood disintegrative disorder (CDD). The third group included those with non-spectrum disorders (n = 37) and with mostly language impairment, mental retardation, and ADHD.

Data analysis was done on those with and without language. Those with language were further sub-divided into four distinct groups (representing the different age groups in the sample). The test authors noted that the children without language had the highest SCQ scores and that those with language had moderate SCQ scores that did not differ among much for the different age groups (2 - 4 years, 5 - 6 years, 7 -10 years or 11 and older). Children with no language, those aged 2 - 4, and 5 - 6 years all had alpha indexes of internal consistency of 0.84 while those aged 7 - 10 or 11 and older had .89 and .93, respectively. The authors further categorized the psychometric properties of the SCQ using the diagnostic classification and found

an alpha index of internal consistency of .81 for autism, .86 for ASD and .92 for non-spectrum diagnoses.

The SCQ has been studied in other cultures. The SCQ has been used in studies with differing levels of specificity and sensitivity found, thus the adjustment of cut-off points was done to match local populations. For instance, in a study with the Greek version of the SCQ among a sample of 130 children aged between seven to ten years, it was found that the SCQ had a specificity and sensitivity of .99 and .96 respectively (Zarokanellou, Kolaitis, Vlassopoulos, & Papanikolaou, 2017). The cut-off points for the ASD and non-ASD group was 15 when assessed using the rater operator characteristic (ROC). To distinguish between ASD and the high-functioning ASD group, the cut-off point was adjusted to 11.

Another cultural study on the SCQ was conducted by Gau and others (2011). The researchers administered a Chinese version of the SCQ to 736 caregivers of children aged two to 18 years. The 736 study participants all had children clinically diagnosed with Autistic Disorder, Asperger's Disorder and PDD-NOS according to the DSM-IV diagnostic criteria. All study caregivers filled out the Chinese version of the SCQ. Among the 736 participants, 317 were also interviewed using the ADI-R to serve as a measure to evaluate the concurrent validity of the Chinese version of the SCQ. An exploratory factor analysis revealed a three-factor structure, i.e., social interaction, repetitive behaviors, and communication. The study authors found test-retest reliability of .77 - .78 of intra-class correlations, and an internal alpha estimate of consistency ranging from .73 to .91 for the three subscales. Correlation between the Chinese SCQ and ADI-R subscales ranged between high to moderate. The SCQ social interaction subscale correlated highly (i.e., .63) with the reciprocal social interaction, and communication (.54 and .65 for verbal and non-verbal communication respectively) but moderately with the stereotyped behavior

subscale (i.e., .21) of the ADI-R subscales. The SCQ repetitive behavior subscale had a correlation of .49 and .36 for stereotyped behaviors and reciprocal social interactions respectively as well as .39 and .31 for communication (verbal and non-verbal communication respectively) on the ADI-R subscales. The SCQ social communication subscale correlated at .56 and .32 (verbal and non-verbal respectively), .37 (stereotyped behaviors) and .31 (reciprocal social interaction) with the ADI-R subscales. Gau and colleagues (2011) found a sensitivity and specificity of .71 and .77 respectively using the authors recommended cut-score of 15. However, when they utilized a cut-score of 12, the sensitivity increased to 0.86 while specificity fell to .60. The researchers concluded that the Chinese version of the SCQ was useful in assessing for the broader autism spectrum (now collectively called ASD under DSM-5) but did not distinguish between DSM-IV autistic disorder and Asperger's Disorder within the spectrum.

In summary, most researchers contend that the SCQ is useful in identifying those at risk of autism, but those interested in using it should consider adjusting its cut-off scores depending on the age and purpose as well as using it in combination with another autism measure (Bölte, Poustka, & Constantino, 2008; Corsello et al., 2007; Gau et al., 2011; Zarokanellou et al., 2017). It has been validated in more than 11 studies (mostly with White culture) and found to have good clinical utility. However, the need for adjustments in the differential cut-off scores in different contexts as well as the non-optimal utility with children aged 2 - 3 years old (Norris & Lecavalier, 2010). Limited studies exist in the application of the SCQ to other cultures and none to the culture targeted for this study.

Social Responsiveness Scale, Second Edition (SRS-2)

The SRS-2 was designed to be a rapid measure of autistic symptoms. The instrument is useful in the identification of mild changes in autistic behaviors among individuals resulting

from interventions targeting the reduction of certain behaviors unique to autism. The SRS-2 has three different forms, i.e., Preschool, School Age and Adult Forms. The instrument assesses social or interpersonal behaviors, communication, and repetitive behavior characteristic of individuals with ASD (Constantino & Gruber, 2012).

The instrument has a total of 65 Likert type items with a scale ranging from 1 (not true), 2 (sometimes true), 3 (often true), and 4 (almost always true). A person familiar with the individual with autism fills out the assessment according to the frequency of the behavior related to autism. Of the 65 items, 35 relate to the social impairment criteria as elaborated in the DSM-IV, i.e., the extent to which a child recognizes social cues, their capacity to interpret and appropriately respond to them as well as their general tendency to socially engage with others. Twenty items relate to the stereotyped or restricted range of interests and six to language deficit assessment. The remaining four items relate to different symptoms associated with autism, but also often observed in other neurodevelopmental or psychological disorders (Constantino & Gruber, 2012).

The assessment takes 15 - 20 minutes to administer. Total raw scores (ranging from 0 to 195) are generated indicating the various levels of severity index for social deficits in a person suspected to have ASD. The total raw scores are converted into *T*-scores and utilized to make clinical or educational decisions depending on whether the individual is severely or moderately affected in their daily social interactions.

There are four levels of severity in the SRS-2. A *T*-score of 59 or below is considered within clinically normal limits. *T*-scores of 60 - 65 are indicative of mild to moderate limitation in activities of daily social interactions. In the DSM-IV-TR, children in this category were mild cases of ASD, PDD-NOS or Asperger's syndrome. *T*-scores of 66 - 75 are associated with

children with moderate deficits in their social interactions, i.e., have substantial interference with their daily social interactions. Children in this category were those having moderate severity of ASD. This falls into Autistic Disorder, PDD-NOS, Asperger's Disorder in the DSM-IV and the Social Communication Disorder and ASD in the current DSM-5. Finally, those with *T*-scores of 76 or higher displayed deficiencies in reciprocal social behavior, which cause a severe disturbance in daily social interactions.

According to the authors, the SRS-2 is not as a standalone assessment in ASD diagnosis. It ought not to replace the extensive use of multiple instruments or professionals' knowledge. The SRS-2 *T*-scores are not "either/or" points for autism presence but a guide for comprehensive ASD assessment. A standard error of measurement of 2 *T*-score points is advisable when making clinical or educational decisions to cater to assessment error in different contextual settings.

The SRS-2 was normed on preschool children (2.6 - 4.6 years), school-aged children (4 - 18 years) and adults (19 years and above). Parents or daycare attendants rate the preschool forms while parents and teachers rate the school-age form. The School Age form standardization caters for normative data of children aged 4 -18 years (to be utilized for this study). The study included a total of 2,025 reports. Parents/caregivers had 1,014 reports and teachers had 1,011 ratings. The sample consisted of 493 (48.7%) males and 518 (51.2%) females recruited from 20 sites in 13 states representative of all the four U.S Census Bureau regions (East, West, Midwest, and South). There were 23.2%, 16.2%, 15.9%, and 44.7% children for the West, East, Midwest, and South regions respectively. Their ethnic/racial background percentages were 5.7% (Asians), 15.8 % (Black/African American), 16.6% (Hispanic/Latino), 0.3% (Native American), 59.5% (White) and 1.6% for others. The norming data provided for parents' educational levels in terms of

percentages were 13.7% (less than high school graduate), 26.1% (high school graduate), 24.8% (some college), and 35.4% for four years of college or more.

The alpha internal consistency for the SRS-2 School Age Form was .95 and .96 for parents and teacher reports, respectively. The inter-rater reliability between parents and teachers was r = .61. Constantino and Gruber (2012) recommended the cut-scores of 70 and 85 to identify ASD conditions in unselected general and clinical populations, respectively. According to them, for example, a cut-score of 70 yielded a sensitivity of .78 and specificity of .94. They also specified that 93% of all children above the raw score of 70 were found to have a diagnosis of ASD on follow-up clinical assessment.

Several studies were identified in the use of the SRS in other cultural contexts. In one study conducted in Tehran (Iran), the Farsi version of the SRS-2 was administered to parents and teachers of 533 children (191 boys and 342 girls) aged 7-11 years old (Tehrani-Doost et al., 2018). All study children had had no record of previous developmental or psychiatric disorder. The SRS-2 and SCQ were administered to parents and teachers. The authors reported a concurrent validity of .43 between the SRS-2 and SCQ. The total scale and subscales of "social communication and behavior" and "restricted repetitive behavior and interests" correlated highly at .99 and .90 respectively supporting the DSM-5 criteria for ASD. Tehrani and colleagues concluded that the Farsi version of SRS was a valid measure in helping to identify children with difficulties in communication and social responsiveness within the community.

Cen and colleagues (2017) investigated the psychometric properties of the Mandarin version of the SRS in Mainland China using a sample of 749 children aged 4 -14 years. The sample included parents of 411 non-ASD children and 338 children who were clinically diagnosed with ASD. An internal consistency between .87 - .92 was found for the total scale

while its test-retest reliability was 0.81 - 0.94. According to the authors, the ROC data for the Mandarin version of the SRS reported that it accurately classifies 69.2 – 97.2% of youth with ASD. The study indicated that the parent-report SRS was a reliable and valid ASD screening instrument that can be used in China, but separate cut-scores for screening and clinical purposes had to be established. They suggest that a cut-score of 56.5 with a sensitivity of .95 specificity of .90 is effective in discriminating between children with and without ASD (Cen et al., 2017).

In another study, a German version of the SRS was administered to 1) parents (mothers and fathers) of 838 non-ASD children, and 2) 527 with ASD and other psychiatric diagnoses, i.e., ADHD, conduct disorders, personality disorders, mental retardation, enuresis, and others. An internal consistency of .91 - .97, a test-retest reliability of .84 - .97 and interrater reliability of .76 and .95 was found (Bölte, Poustka, & Constantino, 2008). A ROC analysis of ASD versus other clinical diagnoses yielded an area under curve [AUC (Area Under the ROC Curve)] of .83 with a total score of 85 having a sensitivity of .73 and specificity of .81.

The SRS is easy to administer and has been validated for use in many countries with good psychometric properties. However, cutoff points are not uniform in other populations hence the need for more multicultural validation studies in places like Africa where the SRS has not been validated.

Childhood Autism Rating Scale, Second Edition (CARS-2)

The Childhood Autism Rating Scale (CARS) was initially called the Childhood Psychosis Rating Scale. It was created following Kanner's (1943) and Creak's (1964) criteria of autism (Schopler, Reichler, DeVellis, & Daly, 1980). The scoring of the instrument is based on observations by experienced raters together with a combination of clinical examination, reports

by parents or other primary caregivers of a child being assessed and psychological tests (Chlebowski, Green, Barton, & Fein, 2010).

The CARS or CARS-2 is used for clinical or educational decisions making as well as assessing change in autism symptomology resulting from rehabilitation efforts. It assesses the presence and severity of autism symptoms and is scored on a Likert scale ranging from 1 (within normal limits for one's age), 2 (mildly abnormal), 3 (moderately abnormal) and 4 (severely abnormal) (Schopler et al., 1980). The CARS-2 had good internal consistency in the ranges of .94 and inter-rater reliability of .71 (Schopler et al., 1980). The total composite scores possible range from 15 to 60. A person with a total score of less than 30 is not considered to have autism while those with total scores above 30 are considered to have autism.

The original CARS was criticized for producing many false positives and over-diagnosing children of two years of age with autism (Chlebowski et al., 2010). Chlebowski and others (2010) investigated the use of CARS in a large sample of toddlers and preschoolers aged two and four years for evaluation of ASD. They found a Cronbach alpha of .91 and proposed increasing the cut-off point for two-year-old from 30 to 32 while retaining the cut-off point of 30 as suggested by Schopler and others (1980), thus increasing the test's specificity and sensitivity. The instrument has been modified to include the standard version (CARS2 –ST) as well as catering for those with high functioning autism (CARS2-HF; Vaughan, 2011).

The standard version (CARS2 –ST) is administered to children under the age six or those above age six, but with significantly impaired communication skills or an IQ of 79 and below. It has a total of 15 items that assess for behaviors characteristic of individuals with autism. The completion of ratings for the CARS2-ST can be done based on information from a single interview or direct observation. CARS2-ST was normed on a total of 1,034 people (78% males

and 22% females) with ASD. The participants' ages ranged from 2 - 36 and were from all the four geographical regions of the U.S Census Bureau. The CARS2-ST total score for internal consistency was .93.

The High Functioning version (CARS2-HF) was designed to assess individuals who are high functioning and suspected to be at-risk for ASD. It has a total of 15 items related to autism symptomology. To assign ratings on the CARS2-HF, the rater must have information from more than one source. This can, for example, be information from direct observation and interviews. Also, important to note is that in the assignment of ratings, the assessed individual is compared to non-ASD peers of the same age with values ranging from 1 - 4 for all the items.

To increase sensitivity across sections, the authors provide for midpoint scores across each rating from 1 - 4 (Vaughan, 2011). It is, therefore, permissible to give midpoint scores like 1.5, 2.5 or 3.5 if raters believe it depicts the best estimate of the behavior being assessed. The CARS2-HF was normed on 994 people with 78% of them being male and 22% females. Their age ranged between 6 and 57 years and represented all the four regions of the U.S Census Bureau (West, Midwest, Northeast, and South). The internal consistency estimate for the CARS-2-HF was 0.96 while inter-rater reliability derived from 239 samples was .95. The CARS-2-HF total score and relationship to the presence of ASD (sensitivity) was .80 and a specificity of .87.

Studies have been conducted to investigate the application of the CARS-2 in other cultures. In one study conducted in Tanzania, 41 children aged 2 - 14 years were assessed for autism using the CARS-2 (Harrison, Zimak, Sheinkopf, Manji, & Morrow, 2014). Thirty of the children were found to have autism while 11 received a diagnosis of global delay. Children suspected to have Fragile X-syndrome (n = 5), cerebral palsy (n = 1), downs syndrome (n = 1), intellectual disability (n = 5) and language delay (n = 1) but could not be confirmed because of

the lack of appropriate assessment instruments constituted the group with global delay. A clinical psychologist synthesized information from the various assessments (parent report, observation during play using the CARS-2 and adaptive function) before providing a final diagnostic decision following the DSM-IV or DSM 5 criteria. A univariate ANOVA using CARS-2 showed that children with ASD had significantly higher scores then their peers with general delay. The average rating for children with ASD on the CARS-2 was 37.75 (*SD* = 5.75) while that for those with general delay was 27.15 (*SD* = 6.79). Although not a validation study, Harrison, colleagues (2017) showed that it is possible to validate autism screening and diagnostic instruments in Africa.

In a validation study in Lebanon, the CARS2-ST Lebanese version was found to have an internal consistency of .99, interrater reliability of .99 and test-retest reliability (conducted after a month from baseline assessment) of .89. This was based on a total study sample of 109 children (*n* = 90 with ASD and *n* = 19 with intellectual disability) aged 4 - 9 years (Akoury-Dirani, Alameddine, & Salamoun, 2013). The researchers in this study excluded item 14 (Level and Consistency of Intellectual Response) because it loaded on a factor of its own and never correlated with other items in the Lebanese version. This Lebanese version of the CARS2-ST therefore consists of 14 items and not 15 as in the original assessment.

Gilliam Autism Rating Scale (GARS-3)

Gilliam developed the Gilliam Autism Rating Scale (GARS-3) to assess children and young adults aged three to 22 years for autism (Karren, 2017). The measure is administered to parents, teachers or other people who spend considerable time with the person being assessed. The GARS-3 was developed following the DSM-IV as well as criteria set out by the Autism Society of America (Lecavalier, 2005). The GARS was criticized for limited sensitivity and

inter-rater reliability when compared to what was reported by Gilliam (Lecavalier, 2005; South et al., 2002).

The GARS-3 was normed based on ratings from parents, teachers, educational diagnosticians, psychologists, speech clinicians, and teacher assistants. A total of 1,859 children and young adults diagnosed with autism, aged three to 22 years, and living in the U.S. formed the study sample for the GARS-3 (Karren, 2017). The instrument can be administered as a questionnaire or as a structured interview. It is composed of up six subscales with a total of 58 items following the DSM 5 definition of ASD (APA, 2013). The subscales include restricted, repetitive behaviors (13 items), emotional responses (8 items), maladaptive speech (7 items), social interaction (14 items), social communication (9 items), and cognitive style (7 items). Rating of autism behaviors is on a Likert type scale ranging from 0 (not at all like the individual), 1 (not much like the individual), 2 (somewhat like the individual) to 3 (very much like the individual; Karren, 2017).

All items of the measure contribute to the final composite score for the scales called the Autism Index. The Autism Index is utilized to determine the presence or absence of autism. The test developers recommend that children who have selective mutism not be assessed using the maladaptive speech and cognitive style subscales, but with the other four remaining sub-scales (Karren, 2017). The scores in the four or six subscales are totaled (depending on whether a child has been rated on four scales because they have selective mutism or all the six subscales for those without selective mutism), then converted into percentiles indicating those equal to or below the person being rated in the normative sample. The percentiles are then scaled to give the Autism Index with a mean of 100 and a standard deviation of 15. Higher Autism Index scores are

indicative of severe autism symptoms. The overall internal consistency across all the age groups was .94 and .93 for the four and six subscales respectively (Gilliam, 2014).

Studies have been conducted to investigate the use of the GARS in other cultures. Caregivers of children aged 3 - 16 years in South Texas were utilized for the validation of the Spanish version of the GARS-2 (Jackson, Little, & Akin-Little, 2013). Parents of a total of 100 children composed of 77 with ASD and 23 without ASD (ADHD n = 9, mental retardation n = 2, other diagnosis n = 3 or no diagnosis n = 9). Out of the entire study sample, 92% of participants reported their ethnicity to be Mexican or Mexican American with the remaining eight percent being "other". A coefficient alpha of .96 was obtained for internal consistency among all items of the GARS-2, Spanish Version. Correlations among the three scales ranged from .81 to .83. It was also found that participants with ASD had significantly higher scores than their peers in the other diagnostic groups. Test-retest reliability as assessed using Pearson's Coefficient was r = .98 although this entailed analysis of data from only 10 of the 100 participants who responded to the call to participate in the re-test study.

23 Question Questionnaire

The 23 Question Questionnaire (23Q) was developed in the context of the African culture (Kakooza-Mwesige et al., 2014). It is a 23-item scale that is used to assess neurodevelopmental disorders for children aged 2 - 9 years. The 23Q includes all ten items of the original Ten Questions (TQ) - an instrument utilized to assess 22,125 children aged 2 - 9 years for childhood disability in Jamaica, Pakistan, and Bangladesh (Durkin, 1995).

Kakooza-Mwesige and colleagues (2014) in collaboration with a Technical Advisory Group (TAG), which included Ugandan and American clinicians as well as epidemiologists, modified the original TQ developed by Durkin (1995) into 23 (yes/no) response items. They

utilized it to assess 1,169 children in Uganda aged 2 - 9 years for neurodevelopmental disabilities. They specifically included an additional five items for all children aged 2 - 9 years, five items for children less than five years of age and an additional three items assessing for visual, hearing or seizure impairments. They translated and back-translated the entire tool into Luganda - a local language used by most people in the study area. If a child failed one or more items on the instrument, he or she was considered positive for potential neurodevelopmental disorders and referred for further clinical evaluation by specialists. Out of a total of 320 children who screened positive on the 23Q, only eight children had ASD after further clinical evaluation by experts. Kakooza-Mwesige and colleagues reported a sensitivity ranging from .55 to .80 for the 23Q.

Justification for Using the SRS and SCQ

There is not one good screening or diagnostic tool that has been psychometrically validated for use in Uganda and Africa at large. The few validation studies have demonstrated promising results in identifying children with autism (Harrison et al., 2014; Kakooza-Mwesige et al., 2014; Smith et al., 2017). The ADOS and ADI that are considered the gold standard in ASD diagnosis take a lot of time to administer and require extensive training for one to utilize. Professionals trained to this level are few and stretched by the high number of patients in Uganda. The SRS and SCQ were chosen for use in this study because they are administered directly to the parents or other primary caregivers who spend considerable time with the child. Both do not require trained raters and yet have strong reliability and validity with the DSM-IV and ICD-10 algorithms for autism criteria. Kakooza-Mwesige and colleagues (2014) were able to show that screening positive on both the additional ASD items in the 23Q and original TQ items increased the sensitivity for receiving a diagnosis of ASD from .49 (with TQ items only) to .69

(with both TQ and ASD items). Further evaluation of the five autism specific items on the 23Q is needed given that it improved the sensitivity of the TQ.

Summary

Assessing for autism takes extensive training and experience working with an autism population to be competent to make a diagnosis (Volker & Lopata, 2008). Autism assessment requires collaborative effort among professionals (Witwer & Lecavalier, 2007). Even with such training and experience, one is still expected to consider information from other professionals before making a diagnosis. There are few practitioners trained to such levels in Uganda. Children with ASD must wait long before a formal diagnosis, or at the worst, go undiagnosed. The situation is worse for children in rural settings because most professionals who are knowledgeable in ASD live and work in urban areas (Bakare & Munir, 2011; Gona et al., 2015). Such extensive assessments like the ADOS-2 and ADI-R considered the gold standard in ASD are a preserve for Ugandan children who can make it to the health facilities in urban settings. Even then their utility in settings where professionals meant to make diagnostic decisions are overwhelmed by the large patient to professional ratio is little. Having instruments that do not require advanced level training can enable the few well-trained professionals to focus their efforts on individuals with a higher likelihood of having ASD. Furthermore, screening tools such as the SCQ, SRS or the ten ASD specific questions on the 23Q can be utilized by less trained professionals often running rural health centers as well. This reduces the long waits parents, especially in low- and medium-income countries, who are often subjected to the waitlist of as long as six months or more (Bello-Mojeed et al., 2017).

CHAPTER 3: METHODOLOGY

Purpose of the Study

The purpose of this study was to investigate the psychometric properties (i.e., internal consistency reliability, concurrent validity, and clinical discriminant validity) of the Social Responsiveness Scale, Second Edition (SRS-2), Social Communication Questionnaire (SCQ), and select 23Q items when used to assess for ASD among children ages four to18 years old in Uganda. This chapter lays out the six research questions, their associated hypotheses, study participants, procedures, and instruments. Prior to field testing, a systematic procedure describes how items of the SRS-2 and SCQ were reviewed and modified based on a small group of parents' and practitioners' input on the original versions. Finally, details concerning the training of research assistants, data management, and data analyses are presented.

Research Questions and Hypotheses

Research Question 1

In the Ugandan context, do total scores for the three ASD instruments (i.e., SCQ, SRS-2, and 23Q) yield adequate internal consistency relative to values reported in the test manual or available studies for those with ASD?

Hypothesis 1a. Within an ASD sample in the Ugandan context (as determined by the DSM-5 diagnostic criteria), the internal consistency (i.e., Cronbach's alpha) for the total score of the SCQ was anticipated to be $\geq .80$.

Hypothesis 1b. Within an ASD sample in the Ugandan context (as determined by the DSM-5 diagnostic criteria), the internal consistency (i.e., Cronbach's alpha) for the total score of the SRS-2 was anticipated to be $\geq .90$.

Hypothesis 1c. Since no existing data were available for the internal consistency of a composite made up of the ASD-related items from the 23Q, no specific reliability value or range of values was/were predicted for this hypothesis. However, it was anticipated that the internal consistency for a 23Q composite for cases with ASD would be lower than that of the SRS-2 and SCQ due to: a) small number of ASD-related items from the 23Q for ASD screening and b) ASD screening had typically been done at the item level for the 23Q.

Research Question 2

In the Ugandan context, do total scores for the three ASD instruments (i.e., SCQ, SRS-2, and 23Q) yield adequate internal consistency relative to values reported in the test manual or available studies for non-ASD cases?

Hypothesis 2a. For a non-ASD sample in the Ugandan context, the internal consistency for the total score of the SCQ was hypothesized to be $\geq$ .80 (i.e., Cronbach's alpha).

Hypothesis 2b. Within a non-ASD sample in the Ugandan context, the internal consistency for the total score of the SRS-2 was predicted to be $\geq$ .90 (i.e., Cronbach's alpha).

Hypothesis 2c. Since no existing data was available for the internal consistency of a composite made up of the ASD-related items from the 23Q, no specific reliability estimate was made for this hypothesis. It was, however, anticipated that the internal consistency for the 23Q composite for non-ASD cases would be lower than that of the SRS-2 and SCQ due to: a) small number of ASD-related items from the 23Q for ASD screening, and b) ASD screening has typically been done at the item level for the 23Q.

Research Question 3

In the Ugandan context, do total scores from the three instruments (i.e., SCQ, SRS-2, and

23Q) yield substantial mean differences consistent with construct-related (i.e., ASD construct) differentiation between ASD and non-ASD groups?

Hypothesis 3a. For the SCQ total score in the Ugandan context, it was anticipated that the mean of the ASD group would be significantly and substantially higher than the mean of the non-ASD group.

Hypothesis 3b. For the SRS-2 total score in the Uganda context, it was anticipated that the mean of the ASD group would be significantly and substantially higher than the mean of the non-ASD group.

Hypothesis 3c. For the 23Q total score of the ASD-related items, it was anticipated that the mean of the ASD group would be significantly and substantially higher than the mean of the non-ASD group.

Research Question 4

In the Ugandan context, will use of the recommended screening cut-scores for each of the three instruments (i.e., SCQ, SRS-2, and 23Q) result in adequate sensitivity and specificity using DSM-5 ASD diagnosis vs. non-ASD cases as the outcome variable?

Hypothesis 4a. Within the Ugandan context, a classification analysis comparing ASD vs. non-ASD based on the SCQ cut-score reported in the manual and the DSM-5 ASD vs. non-ASD diagnostic outcome would result in a significant Chi-square value, as well as sensitivity of > .80 and specificity of > .80.

Hypothesis 4b. Within the Ugandan context, a classification analysis comparing ASD vs. non-ASD based on the SRS-2 cut-score reported in the manual and the DSM-5 ASD vs. non-ASD diagnostic outcome would result in a significant Chi-square value, as well as sensitivity of > .80 and specificity of > .80.

Hypothesis 4c. Using the screening criterion for the ASD items on the 23Q (i.e., one or more ASD-related items affirmatively endorsed) compared to the DSM-5 ASD vs. non-ASD diagnostic criterion, no specific screening sensitivity or specificity target values were hypothesized because of a lack of clear prior findings in the context of a full classification analyses. However, lower sensitivity and specificity for the 23Q ASD-related items was hypothesized relative to the SCQ and the SRS-2, as these other instruments consist of many more items and likely better represent the ASD construct.

Research Question 5

Beyond the recommended cut-scores for screening reported in each test manual, are there more optimal cut-scores for each of the three instruments (i.e., SCQ, SRS-2, and 23Q) in the Ugandan context?

Research Question 6

Are there significant convergent relationships among the total scores of the three instruments (i.e., SCQ, SRS-2, and 23Q) in the Ugandan context?

Hypothesis 6a. Within the Ugandan context, the correlation across the instruments on the total scores of SRS-2 and SCQ was hypothesized to be significant, positive, and substantive (i.e., $\geq$ .50). This was based on the reported correlations between the SRS-2 total score and SCQ total score ranging from .68 in non-ASD group to .50 for the ASD-specific group in the SRS-2 manual.

Hypothesis 6b. Within the Ugandan context, the correlation between the SRS-2 total score and 23Q summed ASD-related items score was hypothesized to be significant and positive but lower than the correlation between the SRS-2 and the SCQ total scores.

Hypothesis 6c. Within the Ugandan context, the correlation between the SCQ total score and the 23Q summed ASD-related items score would be significant and positive but lower than the correlation between the SRS-2 and the SCQ total scores.

Participants

The researcher collected data from a total of 107 participants. All participants were caregivers, who were defined as a parent or person who spends considerable time with the child providing primary care. Study participants were recruited during routine clinical follow-up visits, from regular and special needs schools as well as autism assessment or rehabilitation centers within Kampala and the surrounding districts. Inclusion criteria of caregivers were as follows: 1) live within Kampala and the surrounding districts; 2) adult caregivers of 20 years and older; 3) proficiency in speaking and comprehension of the English language; 4) spent at least 20 hours per week having regular physical interaction with the child; 5) no cognitive disabilities that would impede their judgment.

The study was composed of two groups. Group 1 included 51 caregivers of children aged between 4 and 18 years with a confirmed diagnosis of ASD. With the help of the school, clinic and hospital administrators, caregivers whose children had a diagnosis of autism were invited to be part of the study. Upon agreement to participate, caregivers were requested to provide formal medical or psychological documentation confirming an autism diagnosis for their child. The research assistant reviewed and consulted with the primary researcher before a decision for inclusion in the study based on the availability of clinical documentation. Only parents with clear autism diagnostic documentation for their child were included in the study. Caregivers of children meeting the inclusion criteria were consented to participate in the study. The children were assented (depending on whether they were above or below seven years of age as per the

Ugandan IRB protocol) before the beginning of the study. Those without proper documentation were referred for further assessment with pediatric psychiatrists or clinical psychologists in Butabika and Mulago national referral hospitals.

Group 2 involved 56 caregivers of typically developing children (without ASD, not identified with or suspected of a developmental disability) within the same age range (i.e., 4 to 18 years). A similar caregiver recruitment procedure as the first group was followed, with the difference being that caregivers were asked about the developmental history of their child to rule out the possibility of including undiagnosed autism or a closely related disorder. Caregivers of children meeting the inclusion criteria were consented to participate in the study. The children were assented (depending on whether they were above or below seven years of age as per the Ugandan IRB protocol) before the beginning of the study. Those identified, as being at risk for undiagnosed autism were excluded and referred for further clinical workups with pediatric psychiatrists or clinical psychologists in Butabika and Mulago national referral hospitals.

Procedure

Review and Modification of the SRS-2 and SCQ

Both the SRS-2 and SCQ were developed based on the U.S. culture. To ensure that the items were applicable to the Ugandan culture in terms of linguistic use, context, and content, a request was made of the test patent holders (i.e., Western Psychological Services [WPS]) to purchase the two instruments for pilot use in Uganda. When approved, both instruments were reviewed by five Ugandan parents of varying education levels (i.e., high school, college, and graduate) and three Ugandan trained psychologists (i.e., one community and two clinical psychologists). They provided feedback on the items of the SRS-2 and SCQ after reviewing the items by communicating via e-mail with the researchers in the U.S. As suggested by the parents

and psychologists, item review and adjustment were conducted with the primary researcher, and two PhD level researchers familiar with both the clinical and research aspects of the SRS-2 and SCQ. One researcher also had extensive research experience in psychological testing and measurement. All unclear items were modified for better comprehension and cultural fit.

A total of seven SRS-2 items (i.e., items 5, 25, 29, 30, 39, 50 and 64) were adjusted. For example, item 5 “Doesn’t recognize when others are trying to take advantage of him or her” was changed to “Doesn’t realize when others try to deceive, trick, or use him or her for their own benefit”. The wording was changed because “take advantage of” was less familiar to the mothers with less education. Also, item 29 “Is regarded by other children as odd or weird” was changed to “Is regarded by other children as odd or unusual”. This was because parents in the pilot phase were more familiar with the word “unusual” as compared to “weird”.

Six SCQ items (i.e., items 2, 3, 5, 6, 11, and 34) were iterated to fit the Ugandan research context. For item 2 “Can you have a to and fro “conversation” with her/him that involves taking turns or building on what you have said”, was changed to “Can you have a back and forth “conversation” with her/him that involves taking turns or building on what you have said”? This decision was made basing on the fact that “to and fro” was an unfamiliar expression to several parents in the pilot study. Reviewers suggested that “to and fro” be replaced with “back and forth.” Item 5 “Has she/he ever got her/his pronouns mixed up (e.g., saying you or she/he for I)”? was changed to “Has she/he used the words you or she/he when she/he intended to say I”? Most reviewers and parents in the pilot study were concerned over the meaning of the word “pronoun” hence suggesting that it be left out and emphasis be placed examples. The modified version of both the SRS-2 and SCQ were then submitted to WPS for final approval to be administered to the parent participants.

Both the SRS-2 and SCQ are filled out by caregivers most familiar with the child or person being evaluated. However, for this study, the administration of both instruments was changed. Research assistants administered the items directly to the study caregivers by reading them out and filling in the caregiver responses. The 23Q administration did not change given that it was set up to be administered to the person most familiar with the child by a trained professional as opposed to having the parent fill it out directly.

Training of the Research Assistants for Instrument Administration

Two research assistants working in Butabika and Mulago National Referral Hospitals and known to the lead researcher through prior research engagement were emailed about the part-time research job opportunity to which they agreed to participate. Study details were shared with them through a meeting organized by the lead researcher while in Uganda. Both research assistants had master's level training in Clinical Psychology from Makerere University in Uganda and administered psychological tests as part of their routine work.

The lead researcher provided two half-day trainings to the research assistants on the administration of study instruments, consent and conducting of outreach. Four hours of training were dedicated to proper administration of the study instruments and methods of successfully initiating, conducting, and closing interviews.

Fidelity Check

On the third day of training, each research assistant administered study instruments to a parent of a non-study child with the primary researcher observing the entire procedure (i.e., how they initiated, conducted, and closed interviews) and assessing them for fidelity. A fidelity checklist was created with input from the primary researcher as well as the major academic advisor to the lead researcher (Appendix G).

Field-testing

When granted, the Institutional Review Board (IRB) clearances for each of the Ugandan sites and study host institution (Michigan State University), the researcher reached out to the appropriate school and hospital managers in Butabika hospital and the Special Children's Trust in person for potential study collaboration. Study details together with a request for permission to utilize the school or hospital, as a study site were communicated with the management. When permitted, the researcher provided the administration with letters inviting parents to be part of the study.

Principals or managers helped identify potential caregivers of children aged four to 18 years within in the school (Special Children's Trust) or the hospital clinics in Butabika National Mental Health hospital. Only information of children whose parents were interested in participating in the study was abstracted from available hospital and school records. The school or hospital staff then made telephone calls to interested parents reminding them of the invitation to be part of the study.

Once identified, potential study participants were informed about the ongoing study and asked if they were interested in participating. If they were, a parent together with his/her child (if above seven years) were requested to meet with the researcher for further details about the study. The child was requested to meet with the researcher as per the informed consent guidelines by the Makerere University School of Health Sciences IRB. The research assistants explained the study in its entirety to the potential study participants using simple English and provided an opportunity for them to ask questions before questioning if they were willing to be part of the study. If they agreed to participate in the study, they were consented and/or assented, screened and included in the study.

Data collection took 11 months. The primary researcher together with the two research assistants administered the study instruments to caregivers during the first month of data collection in Uganda. Fidelity checking continued to be used to assess the research assistants' adherence to acceptable interview guidelines for the first month. Both research assistants reached excellent fidelity levels (i.e., 95%) in the administration of instruments and principles of good ethical practice when interacting with research participants. The two research assistants continued with the data collection as the primary researcher returned to the U.S. to continue with his studies. All study caregivers were given an anonymous identification number generated for them and a packet of instruments. A trained research assistant administered all the assessments and demographic information to the caregiver on a one-on-one basis in a secured room within the school, hospital, or other private premises. Instrument administration for each parent took between 35-40 minutes. For all completed protocols, the research assistants double-checked each measure for completeness and accuracy in the presence of the caregiver. If errors were identified, caregivers were requested to clarify.

Measures

Social Communication Questionnaire (SCQ) - Lifetime Form

The SCQ Lifetime Form assesses traits of ASD throughout the lifetime of an individual suspected of having ASD (Lord et al., 1994; Rutter, Bailey, & Lord, 2003). It is a 40-item instrument, with each item scored dichotomously (i.e., Yes/No), and administered to the parent or another caregiver most familiar with the child suspected of having ASD. Self-administration of the screening measure takes approximately ten minutes, and the scoring can be completed in five minutes. Thirty-nine of the 40 items are summed to calculate the total score, which can

range from 0 to 39. Please refer to chapter two for detailed information about the psychometrics of the SCQ.

Social Responsiveness Scale, Second Edition (SRS-2) - School-Age Form

There are three SRS-2 forms, the School-Age (4 – 18 years), Preschool (2.5 and 4.5 years), and the Adult (Self/Relative/Other Report) forms. The SRS-2 School-Age form is a parent or caregiver completed questionnaire, which is used to assess for presence and severity of social impairment of autism spectrum disorder (Constantino & Gruber, 2012). The original SRS (Constantino & Gruber, 2005) and current SRS-2 School-Age form consist of 65 items rated on a Likert scale of 1 to 4 (1 = not true, 2 = sometimes true, 3 = often true, 4 = almost always true). Please refer to chapter two for detailed information about the psychometrics of the SRS-2. Self-administration of the SRS-2 takes between 15 to 20 minutes. For the current study, a research assistant as opposed to them filling it out by themselves administered the SRS-2 to the parents. Its administration took between 25 to 45 minutes. Please refer to chapter two for detailed information about the psychometrics of the SRS-2.

The 23 Question Screener

The 23Q has 23 items used to assess Ugandan children aged 2 - 9 years of age for different neurodevelopmental disabilities (Kakooza-Mwesige et al., 2014). The 23 items have a yes, no, don't know or not applicable responses and a child is considered as a likely positive for neurodevelopmental disability if they fail one or more items. However, it is not used for diagnostic purposes, but is a screener. Children identified as positive are referred for further clinical evaluation. Not much was reported about the psychometrics of the instrument. Kakooza-Mwesige and colleagues reported a sensitivity ranging from 0.55 to 0.80 for the 23Q. However,

for the current study, only ten of the 23Q items specific to ASD screening were used. Please refer to chapter two for detailed information about the psychometrics of the 23Q.

Data Checking and Entry

By the end of the data collection, the research assistants transported all completed instruments to a secure office in Mulago hospital. All hard copies of the instruments were scanned and uploaded to OneDrive (a secure password-protected online platform) for review by the lead researcher. The hard copies were then locked in a secure cabinet at the Global Health Uganda offices in Mulago National Referral hospital.

The research assistants then created password protected database files and sent by e-mail to the lead researcher at Michigan State University and stored through OneDrive, a secured cloud-based tool designed for secure data storage and sharing among approved research collaborators across sites. The lead researcher then entered all the data into the SPSS database system in the study office at Michigan State University, with double data entry and error checklist verification for quality control. In Uganda, the research assistants ensured that all hard copies of study data were securely transported to and stored in locked file cabinets within the Child Health Uganda (CHU) offices in Tororo Main hospital.

Data Analysis

This study provides critical data concerning the reliability, validity, and screening utility of the SCQ, SRS-2, and 23Q ASD-specific items in the Ugandan cultural context. Descriptive statistics for all measure scores were reported, as well as correlations among scores within and between instruments. Internal consistency estimates (i.e., coefficient alpha) for the SCQ and SRS-2 composite scores are reported. Item analyses (e.g., item-total correlations, alpha if item deleted, item discrimination, etc.) were conducted for the SCQ and SRS-2, while only item level

analyses were conducted for the 23Q ASD items, as these items were not typically summed into a composite. The 23Q item responses were recoded as 1 or 0. “Yes” meant the presence of behavior in a child consistent with ASD and was coded as 1, while the other three responses (i.e., “No”, “Not Applicable”, and “Don’t Know”) were coded as 0 indicating absence of behavior.

Discriminant validity of the SCQ and SRS-2 composite scores was examined through: 1) the point-biserial correlation between test score and the DSM-5 dichotomous diagnostic outcome (i.e., ASD vs. non-ASD); 2) predicted mean differences between children with a known ASD diagnosis and those with no known-ASD diagnosis (Non-ASD children); and 3) Receiver Operating Curve (ROC) analyses to determine optimal screening cut-scores for the SCQ and SRS-2 when attempting to differentiate between known ASD and non-ASD children. Screening sensitivity and specificity estimates were reported for various cut-scores for the SCQ and SRS-2, as well as for the number of item endorsements on the 23Q. Optimum screening cut-scores maximized true positive and true negative screening classifications, while minimizing false negative and false positive screening classifications. False negatives were deemed the most problematic screening error, as true positives and false positives go on for further diagnostic evaluation for more accurate differentiation in the real world while true negatives and false negatives both end up getting screened out. All data analyses were performed using SPSS version 25.

CHAPTER 4: RESULTS

This study assessed the psychometric properties of the SRS-2, SCQ and 23Q among children aged 4 -18 years in Uganda. The present chapter reports on the demographics and findings for all six study research questions and associated hypotheses. Research questions 1 and 2 relate to findings on the study instruments' reliability for ASD and non-ASD groups respectively. Research questions 3 and 4 show findings for validity and classification analyses for all the study instruments respectively. Research questions 5 and 6 yield results from the exploration of alternative cut-scores and convergent validity of the total scores for the study instruments.

Demographic Characteristics of Parents

Table 1 provides a detailed breakdown of the parents' demographic information for the ASD and non-ASD groups. When considering both groups together, a total of 102 parents (out of 107) provided demographic information. Five parents had incomplete demographic forms involving several unanswered items. All parents in both groups were of African ethnicity (100%, n = 107). For age, 28 (52.8%) and 27 (55.1%) of non-ASD and ASD parents were between 31 and 40 years of age. For geographic residence, 49 (87.5%) from the non-ASD group and 30 (58.8%) from the ASD group lived in Kampala district. For marital status, 62.5% (n = 35) and 54.9% (n = 28) of the non-ASD and ASD groups were married. Most mothers, (i.e., 89.3% (n = 50) of the non-ASD group and 74.0% (n = 37) of the ASD group) were employed. For fathers, 96.3% (n = 52) of non-ASD and 90.7% (n = 39) of ASD groups were employed. Most parents in the non-ASD (73.9%, n = 34) and ASD (78.6%, n = 22) groups earned a monthly income between 0 (no income) and 3,000,000 Ugandan Shillings (U.S $830). Parents from the non-ASD group mostly had a secondary or polytechnic education (33.9%, n = 19) or a bachelor's degree

(25.0%, *n* = 14). In contrast, most parents of children with ASD had completed ordinary level or 11 years of school (34.0%, *n* = 17) followed by a bachelors, that is, 16 years of school (24.0%, *n* = 12).

Table 1. *Demographic Characteristics of Parents*

		Non-ASD (*n* = 56) *n* (%)	ASD Group (*n* = 51) *n* (%)	Total Sample (*N* = 107) *n* (%)
Parent Age (Years)	≤ 20	--	1 (2.0%)	1 (1.0%)
	21 - 30	8 (15.1%)	11 (22.4%)	19 (18.6%)
	31 - 40	28 (52.8%)	27 (55.1%)	55 (53.9%)
	41 - 50	15 (28.3%)	8 (16.3%)	23 (22.5%)
	51 - 60	2 (3.8%)	2 (4.1%)	4 (3.9%)
District	Kampala	49 (87.5%)	30 (58.8%)	79 (73.1%)
	Luwero	--	1 (2.0%)	1 (0.9%)
	Mbarara	--	1 (2.0%)	1 (0.9%)
	Tororo	--	1 (2.0%)	1 (0.9%)
	Wakiso	7 (12.5%)	18 (35.3%)	25 (23.1%)
Ethnicity	African	56 (100.0%)	51 (100.0%)	107 (100.0%)
Marital Status	Single	8 (14.3%)	8 (15.7%)	16 (15.0%)
	Married	35 (62.5%)	28 (54.9%)	63 (58.9%)
	Living with domestic partner	11 (19.6%)	9 (17.6%)	20 (18.7%)
	Divorced	--	3 (5.9%)	3 (2.8%)
	Widowed	2 (3.6%)	3 (5.9%)	5 (4.7%)
Mother Employed	No	6 (10.7%)	13 (26.0%)	19 (17.9%)
	Yes	50 (89.3%)	37 (74.0%)	87 (82.1%)
Mothers' Occupation	Private Business	10 (17.9%)	22 (43.1%)	32 (29.9%)
	Housewife	5 (8.9%)	5 (9.8%)	10 (9.3%)
	Professional	37 (66.1%)	8 (15.7%)	45 (42.1%)
	Domestic	--	3 (5.9%)	3 (2.8%)
	None	1 (1.8%)	10 (19.6%)	11 (10.3%)
	Others	3 (5.4%)	3 (5.9%)	6 (5.6%)
Father Employed	No	2 (3.7%)	4 (9.3%)	6 (6.2%)
	Yes	52 (96.3%)	39 (90.7%)	91 (93.8%)

Table 1 (Cont'd)

Father's Occupation	None	5 (8.9%)	12 (11.2%)	17 (15.9%)
	Professional	25 (44.6%)	14 (13.1%)	39 (36.4%)
	Private Business	15 (26.8%)	17 (15.9%)	32 (29.9%)
	Others	11 (19.6%)	8 (7.5%)	19 (17.8%)
Income	0 - 3,000,000	34 (73.9%)	22 (78.6%)	56 (75.5%)
	3,001,000 - 6,000,000	10 (21.7%)	4 (14.3%%)	14 (18.9%)
	6,001,000 - 9,000,000	1 (2.2%)	--	1 (1.4%)
	9,001,000 - More	1 (2.2%)	2 (7.1%)	3 (4.1%)
Parent Education **(**Years)	Primary Completed (7 Years)	1 (1.8%)	2 (4.0%)	3 (2.8%)
	Some O' Level (9 Years)	1 (1.8%)	7 (14.0%)	8 (7.5%)
	O' Levels Completed (11 Years)	10 (17.9%)	17 (34.0%)	27 (25.5%)
	A' Level Completed (13 Years)	3 (5.4%)	6 (12.0%)	9 (8.5%)
	Some Post-Secondary or Polytechnic (15 years)	19 (33.9%)	3 (6.0%)	22 (20.8%)
	Bachelors (16 Years)	14 (25.0%)	12 (24.0%)	26 (24.5%)
	Masters or Post-Graduate (18 Years)	8 (14.3%)	3 (6.0%)	11 (10.4%)

Note. Calculations are based on different sample sizes due to missing data for some demographic variables:

[a] District, ethnicity, marital status, mothers' and fathers' occupation were calculated based on a total of 107 parents.

[b] Parental age was based on a total of 102 parents, Mother employed (N = 106), father employed (N = 97), estimated income (N = 74), and education level (N = 106).

Demographic Characteristics of Children of the Participating Parents

Parents provided the demographic information for all 107 study children (n = 51 with ASD, n = 56 non-ASD). A detailed breakdown of their demographic characteristics as reported by parents, are summarized in Table 2. All study children were of African ethnicity with their ages ranging between 4 – 17 years. The children in the non-ASD group were slightly older than those in the ASD group. The age range for the children was 4 -15 years (M = 7.22, SD = 2.46) for the ASD group and 4 -17 years (M = 9.34, SD = 3.52) for the non-ASD group. The average age for ASD first diagnosis was 4.1 years (SD = 1.8) with 80.4% (n = 41) of the ASD group having a comorbidity while 19.6% (n =10) had no known comorbidity. Males diagnosed with ASD were 66.7% (n = 34) with 33.3% (n = 17) of the ASD group being female. Children in the non-ASD group had no known ASD diagnosis and were typically developing. Therefore, demographic information related to age at first diagnosis, comorbidity of ASD or professional who first diagnosed the child was not applicable to the non-ASD group.

Available data showed that medical doctors diagnosed 38.8 % (n = 19), psychiatric clinical officers 24.5% (n = 12), clinical psychologists 18.4% (n = 9), psychiatrists 12.2% (n = 6), other and occupational or speech therapists 6.1% (n = 3) of the children with ASD (Table 2).

Table 2. *Demographic Characteristics of Children*

Demographic Characteristics	ASD (n = 51)	Non-ASD (n = 56)	Total (N = 107)
Child Age in Years	M = 7.22 (SD = 2.46) Range (4-15)	M = 9.34 (SD = 3.52) Range (4-17)	M = 8.34 (SD = 3.24) Range (4-17)
Gender	Males 59.6% (n = 34) Females 34.0% (n = 17)	Males 40.4% (n = 23) Female 66.0% (n = 33)	100.0% (n = 57) 100.0% (n = 50)
Child's Ethnicity			
African	100.0 % (n = 51)	100.0% (n = 56)	100.0% (N = 107)
Age at First Diagnosis	M = 4.1 (SD = 1.8)	N/A	N/A
Comorbidity of ASD	No 19.6% (n = 10) Yes 80.4% (n = 41)	N/A	N/A
Professional who first diagnosed child	Medical Doctor 38.8% (n =19)	N/A	N/A
	Clinical Psychologist 18.4% (n = 9) Psychiatrist 12.2% (n = 6)	N/A	N/A
	Psychiatric Clinical Officer 24.5% (n = 12)	N/A	N/A
	Other & Occupational or Speech Therapist 6.1% (n = 3)	N/A	N/A

Note. Age was missing for 2 cases in the ASD group.

Research Questions and Hypotheses

This section reports the study findings in relation to each of the stated research questions and associated hypotheses. Research questions 1 and 2 related to findings about the reliability of the study instruments. All instruments were assessed for overall reliability among all study participants (i.e., ASD and non-ASD combined) before analyses were conducted within individual groups (i.e., ASD and non-ASD separately). Research question 3 pertained to the

validity of the three study instruments for differentiating between groups known to differ on the intended construct (i.e., ASD and non-ASD). For this hypothesis, differentiation was assessed via mean differences between groups. Continuing with the assessment instrument validity for differentiating between groups, research question 4 examined the screening utility of prescribed cut-scores for each instrument (based on information provided in each instrument's manual or core article[s]). Screening utility was examined in terms of the sensitivity and specificity of each instrument with its recommended cut-score. Research question 5 explored the screening utility of alternative cut-scores for the three instruments in the Ugandan context. While there were no specific directional hypotheses, tentative hypotheses were set based on the range of screening cut-scores using ROC curves. Research question 6 involved examining the convergent validity of the total or composite scores for each instrument via their inter-correlations with each other (i.e., correlations between the instruments). The results of each research question and hypothesis are detailed below.

Research Question 1

In the Ugandan context, do total scores for the three ASD instruments (i.e., SCQ, SRS-2, and 23Q) yield adequate internal consistency relative to values reported in the test manual or available studies for those with ASD?

Hypothesis 1a. *Within an ASD sample in the Ugandan context (as determined by the DSM-5 diagnostic criteria), the internal consistency (i.e., Cronbach's alpha) for the total score of the SCQ was anticipated to be* $\geq$ *.80.*

The Cronbach alpha value obtained from the total scores of the SCQ among the ASD sample was .61. A closer inspection of the reliability statistics showed that if item 4 (inappropriate question) was deleted, the Cronbach alpha would increase to .64, which is not a

significant change in the reliability. Given that the reliability is below the hypothesized value of .80, hypothesis 1a was not supported.

Hypothesis 1b. *Within an ASD sample in the Ugandan context (as determined by the DSM-5 diagnostic criteria), the internal consistency (i.e., Cronbach's alpha) for the total score of the SRS-2 was anticipated to be ≥ .90.*

The Cronbach alpha value obtained from the total score of the SRS-2 in the ASD group (n = 51) was .84. Item analyses revealed that any other item deletion would not substantively improve the reliability of the SRS-2 in this sample. Despite an improvement in internal consistency, the estimate was still below the hypothesized value of .90. Therefore, hypothesis 1b was not supported.

Hypothesis 1c. *Since no existing data were available for the internal consistency of a composite made up of the ASD-related items from the 23Q, there were no guiding estimates for this hypothesis. However, it was anticipated that the internal consistency for a 23Q composite for cases with ASD would be lower than that of the SRS-2 and SCQ due to a) small number of ASD-related items from the 23Q for ASD screening and b) ASD screening had typically been done at the item level for the 23Q.*

The 23Q has ten items intended to screen for ASD (i.e., items 14 to 23 in the order they appear on the instrument form). Items 14 to 18 are used to screen for ASD among children of all ages, while 19 to 23 are additional items administered to children below five years of age for the same purpose (i.e., assess for ASD). Therefore, in typical practice, parents whose children are below five years of age respond to a total of 10 items while those with children five years and above respond to only five items. *In the present study, all ten ASD items were administered to all study participants, regardless of age*. An analysis of internal consistency when combined into a

composite (i.e., total score consisting of the sum of all ten ASD items of the 23Q) yielded a reliability of .60 for the ASD group. Item analyses indicated that if items 14 “Difficulty making eye contact” and 15 “Cries or gets upset if routine is not followed” were deleted, the general reliability of this subset of 23Q items improved. Without these items, reliability increased from .60 to .73. Item analyses conducted after items 14 and 15 were removed, indicated that removal of three additional items further improved internal consistency of the 23Q. These three items were 18 “Repeats phrases exactly as said or heard”; 16 “Takes interest playing with other children”; and 17 “Turns to look when name is called”. When these items were deleted, the reliability increased from .73 to .86. Although the retained items were intended for the screening of ASD in children less than five years of age (i.e., 5.7% of the entire study sample with ASD), the retained items resulted in a reliability of .86 for the 23Q among the whole ASD group. Therefore, combining items 19 through 23 into a total score, for cases with ASD in the Ugandan context, yielded an internal consistency reliability within a range considered adequate for screening. When the total score of the five ASD items recommended for administration to children of all ages (i.e., items 14 –18, the ones suggested for deletion by the above analysis) were used to assess all children with ASD, the Cronbach’s alpha obtained was .13.

Given that the hypothesized reliability was made based on a total score of all the ten ASD-specific items (i.e., .60), hypothesis 1c was supported because the 23Q reliability was below that of the SRS-2 and SCQ at .84 and .61, respectively. However, it is also noteworthy that there was evidence that the internal consistency could be substantively improved through item deletions.

Research Question 2

In the Ugandan context, do total scores for the three ASD instruments (i.e., SCQ, SRS-2, and 23Q) yield adequate internal consistency relative to values reported in the test manual or available studies for non-ASD cases?

Hypothesis 2a. *For a non-ASD sample in the Ugandan context, the internal consistency for the total score of the SCQ was hypothesized to be ≥ .80 (i.e., Cronbach's alpha).*

The Cronbach's alpha value obtained using the total score of the 39-item SCQ for the non-ASD children was .72. Given that it is below the hypothesized value of .80, hypothesis 2a was not supported. The SPSS output following calculation of an instrument's reliability yields item analyses that indicate how internal consistency may improve if each item is deleted while all others are retained. A decision to delete item 19 "Best friends" was taken following the inspection of that SPSS output which showed that the overall SCQ reliability would increase (though not substantively). When item 19 was deleted, reliability slightly increased to .73 for the item set, as shown in Appendix A. Further inspection, following the deletion of item 19, showed that deletion of any other items would not yield substantive improvements in reliability. Note that item 19 was deleted only for purposes of this analysis. Thus, it was retained for other analyses used to examine other research questions and hypotheses.

Hypothesis 2b. *Within a non-ASD sample in the Ugandan context, the internal consistency for the total score of the SRS-2 was predicted to be ≥ .90 (i.e., Cronbach's alpha).*

The obtained Cronbach's alpha for the total score of the SRS-2 was .89, a value that is slightly below the hypothesized score. The reliability did not change even with weaker items deleted, as shown in Appendix B, because all individual items were highly correlated and

dropping any item could not substantively improve the instrument's inter-item consistency. Therefore, hypothesis 2b was not supported.

Hypothesis 2c. *Since no existing data was available for the internal consistency of a composite made up of the ASD-related items from the 23Q, no specific reliability estimate was made for this hypothesis. It was, however, anticipated that the internal consistency for the 23Q composite for non-ASD cases would be lower than that of the SRS-2 and SCQ due to: a) small number of ASD-related items from the 23Q for ASD screening, and b) ASD screening has typically been done at the item level for the 23Q.*

The Cronbach alpha value obtained from the total score of the ten ASD items (i.e., items 14 - 23) for the 23Q when administered to all parents of children regardless of age was .83. Five items were specifically for children aged five years and below (items 19-23). The reliability of these five ASD items administered to parents of all children regardless of age (i.e., items 14 - 18) without the additional items of children less than five years of age (i.e., items 19 -23) was low at .21. Therefore, when used without the additional items for children who are less than five years of age, the composite reliability made up of the ASD-related items moved from relatively high to unacceptably low (see Nunnally & Bernstein, 1994 for standards). A substantial number of children were screened out as false negatives if the additional items were not administered. Also, when item 15 "Child cries or gets upset if you do not do particular routines" was deleted, i.e., utilizing a total of nine items, the 23Q reliability increased to .90. Further, when just items 19 - 23 (the five items intended to screen children who are less than five years of age) were utilized, the 23Q yielded a reliability of .99. However, since the reliability of the 23Q summed scores of the ten ASD items (.83), the strict target of this hypothesis, was higher than that of the SCQ at .72 and lower than that of SRS-2 at .89 for the non-ASD children, hypothesis 2c was not

supported. (Yet, the analyses regarding different item subsets were very informative for how the reliability might be improved).

Research Question 3

In the Ugandan context, do total scores from the three instruments (i.e., SCQ, SRS-2, and 23Q) yield substantial mean differences consistent with construct-related (i.e., ASD construct) differentiation between ASD and non-ASD groups?

Hypothesis 3a. *For the SCQ total score in the Ugandan context, it was anticipated that the mean of the ASD group would be significantly and substantially higher than the mean of the non-ASD group.*

Independent samples *t*-tests were conducted to examine the mean differences between the two groups. Assumptions of equal variance were tested using the Levene's test. Results indicated that there was a significant difference in the variability of the SCQ scores across the two groups. Thus, an independent samples *t*-test was conducted with equal variances not assumed. A significant mean difference was found (t [95.252] = 21.765, $p < 0.001$, $d = 4.278$; for ASD group, $M = 22.61$, $SD = 4.315$, $n = 49$; for non-ASD group $M = 5.41$, $SD = 3.701$, $n = 56$), which was consistent with the ASD group scoring substantially higher, on average, than the non-ASD group on the SCQ. The effect size *d* value of 4.278 was considered large according to Cohen's (1988) effect size standards. Overall, hypothesis 3a was supported.

Hypothesis 3b. *For the SRS-2 total score in the Ugandan context, it was anticipated that the mean of the ASD group would be significantly and substantially higher than the mean of the non-ASD group.*

The SRS-2 was standardized on U.S children and therefore application of the U.S. T-scores to Ugandan children might have been inappropriate. There was a need to find out if the

unstandardized and standardized scores (based on U.S norms) equally discriminated between ASD and non-ASD children in this study. Regardless of using raw (unstandardized scores) or U.S standardized score norms in the SRS-2 manual, significant group mean differences existed between the ASD and non-ASD groups on the total SRS-2 scores. For both score types, the ASD group had higher mean scores than those in the non-ASD group. Using unstandardized (raw) scores, the Levene's test was significant. Consequently, a *t*-test for the SRS-2, with equal variances not assumed, resulted in t (93.949) = 24.057, $p < 0.001$, $d = 4.687$ (ASD $M = 111.16$, $SD = 20.233$, $n = 51$; non-ASD $M = 26.24$, $SD = 15.616$, $n = 55$). Using standardized T-scores, based on US norms--as provided for in the manual, resulted in a significant and large--standardized effect size (t [97.993] = 16.468, $p < 0.001$, $d = 3.214$; ASD $M = 77.88$ $SD = 8.118$, $n = 50$; non-ASD $M = 53.24$, $SD = 7.118$, $n = 55$). Despite the 1.47 reduction in effect size d when using US standardized scores compared to raw scores, the difference between the groups in both cases was substantial. Based on Cohen's (1988) standard descriptions of effect size *d*, the estimate for the SCQ-2 total scores was considered large regardless of whether raw (based on the mean of non-ASD study children) or U.S standardized scores (derived from the manual) were used. Therefore, hypothesis 3b was supported.

Hypothesis 3c. *For the 23Q total score of the ASD-related items, it was anticipated that the mean of the ASD group would be significantly and substantially higher than the mean of the non-ASD group.*

Significant group mean differences existed between the two groups on the total scores for the ten ASD items of the 23Q with the ASD group having a higher mean score compared to the non-ASD group. The Levene's test indicated significant differences in the variances for the two groups. Thus, the *t*-test was adjusted for this issue. A *t*-test for the 23Q with equal variances not

assumed yielded t (103.216) = 3.009, p = .003, d = 0.578 (for ASD group M = 4.06, SD = 1.943, n = 51; for non-ASD M = 2.95, SD = 1.872, n = 56). The mean difference was significant, and the effect size was moderate in magnitude. Also, when analyzed using only the five ASD items intended for children less than five years of age (i.e., items 19 - 23), the results yielded a significant and large mean difference (t [75.529] = 5.287, $p < 0.001$, d = 1.049, equal variances not assumed; ASD M = 3.08, SD = 1.036, n = 51; non-ASD M = 2.21, SD = .563, n = 56) in the expected direction. Overall, hypothesis 3c was specifically aimed at the 10-item set of 23Q ASD items and was supported.

Research Question 4

In the Ugandan context, will use of the recommended screening cut-scores for each of the three instruments (i.e., SCQ, SRS-2, and 23Q) result in adequate sensitivity and specificity using DSM-5 ASD diagnosis vs. non-ASD cases as the outcome variable?

For sensitivity, the SCQ identified almost all of children with ASD (98%) using the recommended cut-score (i.e., 15) and screened out 94.6% of the non-ASD children (specificity). The SRS-2, with the raw cut-scores of 85 and 70 (recommended for clinical and research samples respectively) identified 84.0% and 98.0% of cases with ASD (sensitivity) respectively, and 100.0% and 96.4% of cases without ASD (specificity) respectively. SRS-2 screening results differed when the cut-scores were applied in the standardized T-score metric (see results for both raw and standardized scores under hypothesis 4b). Endorsing one or more items of the 10 ASD specific items on the 23Q identified all children with ASD (i.e., sensitivity of 100.0%), but had a low specificity, i.e., correctly identified none of study children without ASD, respectively. Below, amidst addressing the specific hypotheses, is a detailed breakdown for classification analyses and screening utility of the SCQ, SRS-2, and 23Q.

Hypothesis 4a. *Within the Ugandan context, a classification analysis comparing ASD vs. non-ASD based on the SCQ cut-score reported in the manual and the DSM-5 ASD vs. non-ASD diagnostic outcome would result in a significant Chi-square value, as well as sensitivity of > .80 and specificity of > .80.*

According to Lord and others (1994), an SCQ cut-off score of 15 or more is indicative of a strong possibility for a clinical diagnosis of ASD. To investigate the applicability of this cut-score among Ugandan children, a classification analysis utilizing a cut-score of 15 or more was conducted on a total of 105 participants (n = 49 with ASD, n = 56 non-ASD). The Chi-square test was used to statistically assess the significance of the relationship between SCQ classifications and actual diagnostic status of the cases. The Chi-square (χ^2 [1, N = 105] = 85.926, p < .001) was statistically significant. The cut-score of 15 yielded a sensitivity of 95.9% and specificity of 94.6% (see Table 3) for children in this study. This implies that using the SCQ cut-score of 15 correctly identified 96% of the sample who had ASD while accurately excluding 95% without ASD. Given that the Chi-square result was significant and that both sensitivity and specificity were higher than predicted (i.e., > .80), hypothesis 4a was supported.

Table 3. *Sensitivity and Specificity of the SCQ Total Raw Score based on Cut-score of 15*

			Research Group		Total
			Non-ASD	ASD	
SCQ Classification (15 Cut-score)	Non-ASD	Count	53	2	55
		% within Research Group	94.6%**	2.0%*	
	ASD	Count	3	47	50
		% within Research Group	6.0%*	95.9%**	
Total		Count	56	49	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the total row and total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false

negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Hypothesis 4b. *Within the Ugandan context, a classification analysis comparing ASD vs. non-ASD based on the SRS-2 cut-score reported in the manual and the DSM-5 ASD vs. non-ASD diagnostic outcome would result in a significant Chi-square value, as well as sensitivity of > .80 and specificity of > .80.*

Constantino and Gruber (2012) recommended 70 and 85 as cut-scores for general research and clinical decision-making purposes respectively. The results of the classification analysis for the raw and standardized cut-score of 70 (based on U.S norms in the manual) are in Tables 4 and 5 respectively. Those for the raw and standardized cut-score of 85 (based on U.S. norms in the manual) are in Tables 6 and 7, respectively. For the SRS-2 raw cut-score of 70, results indicated that the Chi-square was significant (χ^2 [1, N = 105] = 93.361, p < .001, with a continuity-corrected Chi-square value of 89.622 and associated with a p-value < .001, sensitivity of .98 (i.e., 98%), and specificity of .964 (i.e., 96.4%). When applied in the context of norm-referenced (standardized) T-scores, the SRS-2 standardized cut-score of 70 resulted in a significant Chi-square (χ^2 [1, N = 105] = 80.105, p < .001, with a continuity-corrected Chi-square value of 76.588 and associated with a p-value < .001, sensitivity of .86 (i.e., 86%), and specificity of 1.00 (i.e., 100%). These results indicated that the raw SRS-2 cut-score of 70 accurately identified 98.0% (compared to 86.0% identified using the standardized cut-score of 70) of children with ASD, while only 2% of ASD cases were missed and inaccurately screened out (i.e., false negative cases) in comparison to 14.0% falsely screened out as non-ASD using the same standardized score (i.e., norm-referenced T-score of 70). Also, 70 as a raw score was able to accurately screen out 96.4% of the non-ASD cases. Thus, only 3.6% of non-ASD cases were identified as having ASD (i.e., false positive cases). The standardized (norm-referenced T-score)

cut-score of 70, however, accurately screened out all non-ASD children (i.e., had a specificity of 100%).

For the SRS-2 raw (unstandardized) cut-score of 85, results indicated that the Chi-square was significant (χ^2 [1, N = 105] = 77.000, p < .001), with a continuity-corrected Chi-square value of 73.540 and associated p-value < .001, sensitivity of .84 (i.e., 84.0%), and specificity of 1.0 (i.e., 100.0%; Table 6). However, using a norm-referenced standardized T scale cut-score of 85, a significant Chi-square resulted (χ^2 [1, N = 105] = 17.769, p < .001, with a continuity-corrected Chi-square value of 15.429, associated p-value of < .001), sensitivity of .28 and specificity of 1.0 (see Table 7). These results indicated that using a raw (unstandardized) cut-score of 85 accurately identified 84.0% cases with ASD and accurately screened out 100% of non-ASD cases. Thus, the SRS-2 mis-classified 16.0% of children with ASD as non-ASD cases. Despite having a specificity of 1.0, the standardized cut-score of 85 had a very low sensitivity compared to its raw (unstandardized) cut-score (i.e., a false negative rate of .72).

Both the raw (unstandardized) and standardized cut-score of 70 yielded high sensitivity and specificity. However, only the raw (unstandardized) cut-score of 85 yielded a good sensitivity (i.e., .84) in this study. The standardized cut-score of 85 had low sensitivity (i.e., .28). Also, both SRS-2 cut-scores (i.e., 70 and 85) had excellent specificities, classifying close to all non-ASD children as true negatives. Since the SRS-2 is usually scored and interpreted using standardized T-scores, results for this score type were used to evaluate hypothesis 4b. For standardized T-scores, results for the cut-score of 70 (i.e., sensitivity of 98.0 and specificity of 96.4) were consistent with the hypothesis, but results for the cut-score of 85 (i.e., sensitivity of .28 and specificity of 1.0) fell well short of the .80 benchmark for sensitivity. Thus, overall, hypothesis 4b was supported for a cut-score of 70, but not for a cut-score of 85.

Table 4. *Sensitivity and Specificity of the SRS-2 Total Score Based on Raw Cut-score of 70*

			Research Group		Total
			Non-ASD	ASD	
SRS-2	Non-ASD	Count	53	1	54
		% within Research Group	96.4%**	2.0%*	
	ASD	Count	2	49	51
		% within Research Group	3.6%*	98.0%**	
Total		Count	55	50	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the Total row and Total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Table 5. *Sensitivity and Specificity of the SRS-2 Total Standardized Cut-score (T-scores) of 70*

			Research Group		Total
			Non-ASD	ASD	
SRS-2	Non-ASD	Count	53	1	54
		% within Research Group	96.4%**	2.0%*	
	ASD	Count	2	49	51
		% within Research Group	3.6%*	98.0%**	48.5%
Total		Count	55	50	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the Total row and Total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Table 6. *Sensitivity and Specificity of the SRS-2 Total Raw Cut-score of 85*

			Research Group		Total
			Non-ASD	ASD	
SRS-2	Non-ASD	Count	55	8	63
		% within Research Group	100.0%**	16.0%*	
	ASD	Count	0	42	42
		% within Research Group	0.0%*	84.0%**	
Total		Count	55	50	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the total row and total column, the upper left

quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Table 7. *Sensitivity and Specificity of the SRS-2 Total Standardized Cut-score (T-scores) of 85*

			Research Group		Total
			Non-ASD	ASD Group	
SRS-2	Non-ASD	Count	55	36	91
		% within Research Group	100.0%**	72.0%*	
	ASD	Count	0	14	14
		% within Research Group	0.0%*	28.0%**	
Total		Count	55	50	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the Total row and Total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Hypothesis 4c. *Using the screening criterion for the ASD items on the 23Q (i.e., one or more ASD-related items affirmatively endorsed) compared to the DSM-5 ASD vs. non-ASD diagnostic criterion, no specific indices for screening sensitivity or specificity target values were used because of a lack of clear prior findings. However, lower sensitivity and specificity for the 23Q ASD-related items was hypothesized relative to the SCQ and the SRS-2, as these other instruments consist of many more items and likely better represent the ASD construct.*

The Chi-square test was used to assess the relationship for statistical significance between 10-item 23Q dichotomous classifications and known DSM-5 ASD vs. non-ASD classifications. The 2 x 2 classification analysis (i.e., cross-tabulation of ASD vs. non-ASD according to 23Q and ASD vs. non-ASD using DSM-5 diagnosis) is presented in Table 8. The endorsement of one or more of the 10 ASD items of the 23Q, as recommended by the authors, for one to screen positive for a potential ASD diagnosis yielded a sensitivity of 1.0 and specificity of 0.0 among children in the study. The recommended endorsement score of one or more despite identifying

all ASD cases equally had a very high false positive rate. The sensitivity of the 10 summed ASD items for the 23Q was in the same range as that of the SRS-2 and SCQ cut-scores of 70 and 15, respectively (i.e., .98 for both cut-scores) but better than that of the SRS-2 clinical cut-score of 85 (i.e., .82). For specificity, the 23Q ten summed ASD items performed much poorer, identifying none (i.e., 0.0%) of true negative children, when compared to both SRS-2 cut-scores of 70 and 85 (i.e., 96.3% and 100% respectively) and SCQ cut-score of 15 at 94.6% in this study sample. Therefore, given that the 23Q sensitivity was slightly higher than that of the SRS-2 (for both general and clinical cut-scores) and that of the SCQ cut-score of 15, but substantially lower in specificity when compared to both the SRS-2 and SCQ, hypothesis 4c was not supported. However, it should be clear that the 10 ASD 23Q items ended up screening all cases positive for ASD and none of the cases negative for ASD (regardless of actual classification). This suggests very poor overall performance relative to the SCQ and SRS-2 when the endorsement of one or more items is used as the cut-off for the 10 23Q ASD items.

Table 8. *Sensitivity and Specificity of the 23Q Based on Using a Cut-score of 1 or more on the Ten summed ASD Items*

			Research Group		Total
			Non-ASD	ASD	
23Q (≥1 Cut-score)	ASD	Count	56	51	107
		% within Research Group	100.0%*	100.0%**	
Total		Count	56	51	107
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the total row and total column, the upper left quadrant is false positive cases and upper right quadrant is true positive cases (i.e., sensitivity as a percent). Though not explicit in the table, specificity was 0.00% (i.e., all non-ASD misclassified as ASD by the 23Q).

Research Question 5

Beyond the recommended cut-scores for screening reported in each test manual, are there more optimal cut-scores for each of the three instruments (i.e., SCQ, SRS-2, and 23Q) in the Ugandan context?

For this study, the SCQ cut-score of 10 (as compared to 15 recommended by SCQ authors) was found to more optimally distinguish between ASD and non-ASD children. (However, it must be noted that the sensitivity and specificity of the cut-score of 15 were both excellent according to standards from Nunnally and Bernstein [1994].) The cut-score of 10 identified all children at risk of ASD, which was not the case with 15. For the SRS-2, a cut-score of 62 performed better than the author-recommended cut-scores of 70 and 85 (for research and clinical decision making, respectively). Using the 23Q, a total cut-score of 3 performed better than 1 (i.e., the author-recommended cut-score) in simultaneously identifying those at-risk of ASD, while screening out those who do not have ASD. Below is a detailed breakdown of the exploration of alternative cut-scores showing analyses that led to these new cut-scores for the three study instruments.

Exploration of alternative cut-scores for the SCQ. A Receiver Operating Characteristic (ROC) Curve was utilized in assessing for an alternative cut-score yielding a more optimal balance between sensitivity and specificity for the SCQ total score in the study sample (see Table 9). The area under the ROC curve (AUC) quantifies the proportion of accurately classified cases or how well an instrument's cut-score differentiates between groups--such as those with and without ASD for this study. The ROC sensitivity and specificity proportions each range from 0 (accurately identifying none) to 1 (accurately identifying all).

Table 9. *Receiver Operating Characteristic (ROC) Curve of the SCQ Total Cut-Scores*

Possible SCQ Cut-Scores	Sensitivity	Specificity
-1.00	1.000	0.000
.50	1.000	0.071
1.50	1.000	0.161
2.50	1.000	0.232
3.50	1.000	0.304
4.50	1.000	0.411
5.50	1.000	0.518
6.50	1.000	0.679
7.50	1.000	0.768
8.50	1.000	0.839
10.00**	**1.000** $^{++}$	**0.929** $^{+}$
12.00	0.980	0.946
14.00	0.959	0.946^{+}
15.00*	**0.959** $^{++}$	**0.946**$^{+}$
17.50	.857	1.000
18.50	.837	1.000
19.50	.735	1.000
20.50	.673	1.000
21.50	.633	1.000
22.50	.592	1.000
23.50	.531	1.000
24.50	.367	1.000
25.50	.245	1.000
26.50	.204	1.000
27.50	.082	1.000
28.50	.061	1.000
30.00	.020	1.000
32.00	.000	1.000

Notes. Bolded black numbers show approximations of the sensitivity and specificity of different SCQ cut-scores. * Shows an approximation of the author recommended cut-score of 15 (which is the same as the values for a cut-score of 14 depicted in the table) while ** shows the performance of 10 as an alternative cut-score. $^{+}$ and $^{++}$ their respective sensitivity and specificity.

The cut-score of 15 as proposed by the manual had a sensitivity of .96 and specificity of .95 (see Table 10), thus, close to all the children in the study sample were screened out as true negatives or in as true positives. Further analysis of the ROC curve showed that lowering the cut-

score to 10 appropriately distinguished non-ASD children from those with ASD by identifying all ASD children as ASD (i.e., sensitivity of 1.0) although specificity dropped to .93 (see Table 11). With the 49 children with ASD in the study sample, the cut-score of 10 correctly identified all 49 children (a positive predictive value of 1.0). For the non-ASD group, the cut-score of 10 displayed a negative predictive value of .93 (i.e., identified 52 out of 56 children without ASD). The false positive percentage using 10 as a cut-score was low at 7.1% (4 of 56 without ASD) with a false negative rate of 0% (0 out of 49 children with ASD) for this study sample (see Table 11).

Table 10. *Sensitivity and Specificity of the SCQ Based on the Cut-score of 15*

			Research Group		Total
			Non-ASD	ASD	
SCQ Classification (15 Cut-score)	Non-ASD	Count	53	2	55
		% within Research Group	94.6%**	2.0%*	
	ASD	Count	3	47	50
		% within Research Group	6.0%*	95.9%**	
Total		Count	56	49	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the total row and total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Table 11. *Sensitivity and Specificity of the SCQ Based on the Cut-score of 10*

			Research Group		Total
			Non-ASD	ASD	
SCQ Classification (10 Cut-score)	Non-ASD	Count	52	0	52
		% within Research Group	92.9%**	0.0%*	
	ASD	Count	4	49	53
		% within Research Group	7.1%*	100.0%**	
Table 11 Cont'd					
Total		Count	56	49	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the total row and total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Exploration of alternative cut-scores for the SRS-2. A summary of the performance of the cut-scores 70, 85 and 62 when used both as unstandardized (raw) and standardized cut-scores is provided in Table 12. Standardized scores were based on T-scores provided in the SRS-2 manual while raw scores were unstandardized scores of parental rating of study children on the SRS-2, respectively.

Table 12. *Summary of Sensitivity and Specificity of Recommended Versus Alternative SRS-2 Raw and Standardized Cut-Scores*

Cut-Score	Sensitivity	Specificity	False Positive	False Negative
70 (Raw)	98.0%	96.4%	3.6%	2.0%
70 (Standardized)	86.0%	100%	0.0%	14.0%
85 (Raw)	84.0%	100.0%	0.0%	16.0%
85 (Standardized)	28.0%	100.0%	0.0%	72.0%
62 (Raw)	100%	96.4%	3.6%	0.0%
62 (Standardized)	100%	96.4%	3.6%	0.0%

Notes. All standardized scores are based on U.S. T-scores provided in the SRS-2 Manual. Raw scores are based on unstandardized raw scores of study children.

Following a review of the ROC curve coordinates and associated sensitivity and specificity values using raw scores (see Table 13), the cut-score of 62 performed slightly better than the author recommended cut-scores of 70 and 85.

Table 13. *Receiver Operating Characteristic (ROC) Curve of the SRS-2 Alternative Total Cut-Scores Using Raw Scores*

Positive if Greater Than or Equal To[a]	Sensitivity	Specificity
2	1.000	0.000
4	1.000	0.018
7	1.000	0.036
9.5	1.000	0.073
10.5	1.000	0.109
11.5	1.000	0.145
12.5	1.000	0.164
14	1.000	0.218
15.5	1.000	0.236
16.5	1.000	0.255
17.5	1.000	0.273
18.5	1.000	0.327
19.5	1.000	0.364
20.5	1.000	0.400
21.5	1.000	0.436
22.5	1.000	0.491
23.5	1.000	0.509
24.5	1.000	0.564
25.5	1.000	0.600
26.5	1.000	0.636
28	1.000	0.655
29.5	1.000	0.673
30.5	1.000	0.691
32	1.000	0.745
34	1.000	0.782
35.5	1.000	0.800
36.5	1.000	0.818
38.5	1.000	0.836
42.5	1.000	0.873
45.5	1.000	0.891
47	1.000	0.909

Table 13 (Cont'd)

52.5	1.000	0.927
57.5	1.000	0.945
62* **	**1.000**	**0.964**
68.5	0.980	0.964
70*	**0.980**	**0.964**
71.5	0.980	0.982
72.5	0.961	0.982
76	0.941	0.982
79.5	0.922	0.982
80.5	0.902	0.982
82	0.902	1.000
83.5	0.882	1.000
84.5 **	**0.843**	**1.000**
86	0.824	1.000
89	0.804	1.000
93.5	0.784	1.000
97	0.745	1.000
99	0.725	1.000
101.5	0.706	1.000
103.5	0.686	1.000
104.5	0.667	1.000
106	0.647	1.000
107.5	0.627	1.000
109	0.608	1.000
111	0.569	1.000
112.5	0.549	1.000
114.5	0.490	1.000
117.5	0.471	1.000
120	0.412	1.000
121.5	0.392	1.000
122.5	0.373	1.000
124	0.353	1.000
125.5	0.275	1.000
127	0.235	1.000
128.5	0.196	1.000
129.5	0.176	1.000
130.5	0.137	1.000
131.5	0.098	1.000
134	0.078	1.000
136.5	0.059	1.000
138.5	0.039	1.000
147	0.020	1.000

Table 13 (Cont'd)

155	0.000	1.000

Notes. *** shows the performance (sensitivity and specificity) of an alternative cut-score of 62. * and ** depict the sensitivity and specificity of the author recommended research and clinical cut-scores (i.e., 70 [same sensitivity and specificity as 68.5 in the table] and 85 respectively) among study children.

When used to classify children in the study, 62 as a cut-score (raw) classified all children with ASD correctly, that is 1.0 for ASD and .96 for correct classification of non-ASD (see Table 14). This was slightly better than the sensitivity of .98 but the same for the specificity (i.e., .96) obtained when using the raw cut-score of 70. The alternative cut-score of 62 also performed better than the authors' recommended cut-score for clinical decision making (i.e., 85 raw cut-score) which had a sensitivity of .84. However, for specificity, the unstandardized or raw cut-score of 85 identified all non-ASD correctly (i.e., specificity of 1.0) hence performing better than 62 in that respect.

Table 14. *Detailed Breakdown of Sensitivity and Specificity Based on an SRS-2 Raw-score of 62 Classification Analysis*

			Research Group		Total
			Non-ASD	ASD	
Classification with 62 as Alternative Cut-score	Non-ASD	Count	53	0	52
		% within Research Group	96.4%	0.0%	
	ASD	Count	2	50	38
		% within Research Group	3.6%	100.0%	
Total		Count	55	50	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the total row and total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Among standardized T-scores based on U.S. norms provided in the manual, the ROC curve coordinates showed that a T-score of 62 was a better cut-score than standard scores of 70

or 85 (see Table 15). When applied to classify the sample in this study, the cut-score of 62 maintained a high sensitivity (i.e., 1.00) with a slight drop in specificity to .964 (See Table 16).

Table 15. *Receiver Operating Characteristic (ROC) Curve of the SRS-2 Standardized Alternative Total Cut-Scores (U.S. Norms)*

Positive if Greater Than or Equal To[a]	Sensitivity	Specificity
39.00	1.000	0.000
40.50	1.000	0.036
41.50	1.000	0.091
42.50	1.000	0.109
43.50	1.000	0.127
44.50	1.000	0.164
45.50	1.000	0.236
46.50	1.000	0.255
48.00	1.000	0.291
49.50	1.000	0.327
52.00	1.000	0.345
54.50	1.000	0.382
55.50	1.000	0.473
56.50	1.000	0.564
57.50	1.000	0.673
58.50	1.000	0.782
59.50	1.000	0.873
60.50	1.000	0.927
62.00	**1.000**	**0.964**
64.00	0.980	0.964
65.50	0.960	0.982
66.50	0.920	0.982
67.50	0.900	0.982
68.50	0.860	0.982
69.50	0.860	1.000
70.00	**0.860**	**1.000**
71.00	0.780	1.000
72.50	0.720	1.000
73.50	0.680	1.000
74.50	0.600	1.000
75.50	0.540	1.000
76.50	0.500	1.000
77.50	0.480	1.000
78.50	0.420	1.000
79.50	0.400	1.000

Table 15 (Cont'd)

80.50	0.360	1.000
81.50	0.340	1.000
83.00	0.320	1.000
85.00	**0.280**	**1.000**
85.50	0.240	1.000
86.50	0.180	1.000
87.50	0.140	1.000
88.50	0.100	1.000
89.50	0.060	1.000
90.50	0.040	1.000
94.50	0.020	1.000
99.00	0.000	1.000

Note. Scores are in the norm-referenced T-score metric (U.S. norms). Cut-scores were standardized for sex (males and females combined) as rated by a parent.

Table 16. *Sensitivity and Specificity Based on an SRS-2 T-score Classification of 62 in Study Children*

			Research Group		Total
			Non-ASD	ASD	
Classification Using 62 as T –Cut-score	Non-ASD	Count	53	2	55
		% within Research Group	96.4%**	3.6%*	
	ASD	Count	0	50	50
		% within Research Group	0.0%*	100.0%**	
Total		Count	55	50	105
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the Total row and Total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Exploration of alternative cut-scores for the 23Q. In the 23Q screening, a cut-score of 1, as used by Kakooza-Mwesige and colleagues (2014) for identification of children with ASD, failed to adequately discriminate between the children in the different groups (Table 17). Despite identifying all children with ASD, the cut-score of 1 equally screened all non-ASD individuals in

as false positives. A further classification analysis was performed using the cut-scores of 3 (Table 18) and 4 (Table 19). A sensitivity and specificity of .77 and .68 for the cut-score of 3 and .55 and .80 for the cut-score of 4 were attained. Although 4 as a cut-score had a higher specificity than 3, it had a much lower sensitivity. Therefore, a cut-score of 3 appeared to balance sensitivity and specificity more appropriately (compared to cut-scores of both 1 and 4) for the Ugandan children in this sample. Thus, for research question five, each of the three instruments (i.e., SCQ, SRS-2, and 23Q) yielded different, more optimal alternative cut-scores of 10, 62 and 3, respectively.

Table 17. *Receiver Operating Characteristic (ROC) Curve of the 23Q Ten Summed ASD Items' Total Cut-scores*

Positive if Greater Than or Equal To[a]	Sensitivity	Specificity
Table 17 (Cont'd)		
1.0	**1.000**	**0.000**
2.0	.961	0.036
3.0[+]	.765[+]	0.679[+]
4.0	.549	0.804
5.0	.353	0.857
6.0	.196	0.857
7.0	.137	0.857
8.0	.078	0.964
9.0	.020	1.000
10.0	.000	1.000

Note. Bolded number represents the sensitivity and specificity of the cut-score consistent with the authors' recommended score of one while [+] represents the performance of alternative cut-scores.

Table 18. *23Q Sensitivity and Specificity of the 23Q Based on a cut-score of 3.0*

			Research Group		Total
			Non-ASD	ASD	
23Q Classification Using 3.0 Cut-score for ASD	Non-ASD	Count	38	12	50
		% within Research Group	67.9%**	23.5%*	
	ASD	Count	18	39	57
		% within Research Group	32.1%*	76.5%**	
Total		Count	56	51	107
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the total row and total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Table 19. *Sensitivity and Specificity Based on the 23Q Cut-score of 4.0*

			Research Group		Total
			Non-ASD	ASD	
23Q Classification based 4.0 Cut-score	Non-ASD	Count	45	23	68
		% within Research Group	80.4%**	45.1%*	
	ASD	Count	11	28	39
		% within Research Group	19.6%*	54.9%**	
Total		Count	56	51	107
		% within Research Group	100.0%	100.0%	

Note. ** indicates accurate or true screening results and * indicates inaccurate or false screening results. Each column sums to 100%. Ignoring the Total row and Total column, the upper left quadrant is true negative cases (i.e., specificity as a percent), upper right quadrant reflects false negative cases, lower left quadrant is false positive cases, and lower right quadrant is true positive cases (i.e., sensitivity as a percent).

Research Question 6

Are there significant convergent relationships among the total scores of the three instruments (i.e., SCQ, SRS-2, and 23Q) in the Ugandan context?

Research question 6 investigated the correlations between the total scores of the three instruments (i.e., SCQ, SRS-2, and 23Q). Below is a breakdown of the convergent validity

hypotheses and results for each. The total sample (i.e., ASD and non-ASD cased combined) was used to calculate all correlations.

Hypothesis 6a. *Within the Ugandan context, the correlation across the instruments on the total scores of SRS-2 and SCQ was hypothesized to be significant, positive, and substantive (≥.50). This is based on the reported correlations between the SRS-2 total score and SCQ total score ranging from .68 in non-ASD group to .50 for ASD-specific group in the SRS-2 manual.*

The SRS-2 and SCQ total scores had a significant, positive and substantive correlation: r (105) = .93, p < .001. The significance, direction, and general magnitude of the relationship were as predicted. Thus, hypothesis 6a was supported.

Hypothesis 6b. *Within the Ugandan context, the correlation between the SRS-2 total score and 23Q summed ASD-related items score will be significant and positive but lower than the correlation between the SRS-2 and the SCQ total scores.*

There was a non-significant positive correlation r (105) = .19, p = .053 between the SRS-2 total score and the 23Q summed ASD-related items (10 ASD 23Q items, i.e., 23Q items 14 – 23). When only the five 23Q items (i.e., 19 - 23) typically used in screening for ASD in children below five years of age were utilized, the correlation between the five-item 23Q score and the SRS-2 was higher and significant (r [105] = .35, p < .001). When compared to the magnitude of the correlation between the SCQ and SRS-2 total scores (i.e., r = .93), the correlations involving variations of the summed 23Q ten ASD-related items with the SRS-2 total score were both lower (i.e., r = .19 and r = .35, respectively) and only significant when the five ASD items meant to assess children below five years (items 19 – 23) were used. Hypothesis 6b was not supported for the 10-item 23Q summative scale (i.e., items 14 – 23), because the correlation was not significant.

Hypothesis 6c. *Within the Ugandan context, correlation between the SCQ total score and the 23Q summed ASD-related items score will be significant and positive but lower than the correlation between the SRS-2 and the SCQ total scores.*

The correlation between the SCQ total score and 23Q summed 10 ASD-related items was positive but not significant r (105) = .17, p = .082. The correlation magnitude was lower than that between the SRS-2 and SCQ total scores (i.e., r = .93). However, the SCQ total score correlation with the five ASD items used to screen for ASD among children younger than five years (items 19 - 23) was moderate and significant (i.e., r [105] = .36, p < .001) but still having a lower correlation than that between the SRS-2 and SCQ total scores. Hypothesis 6c was not supported when using all the 10 ASD 23Q items (items 14 – 23), because the correlation was not significant.

Findings from all research questions and related hypotheses are summarized below in Table 20. This table is intended to put all of the important results in one place for the reader.

Table 20. *Research Hypotheses Findings Summary*

Hypotheses	Hypotheses Outcome
1a) SCQ reliability for ASD group to be ≥ .80.	Not supported
Table 20 Cont'd	
1b) SRS-2 reliability for ASD group to be ≥ .90.	Not supported
1c) 23Q composite reliability lower than that of the SRS-2 and SCQ for ASD group.	Supported
2a) SCQ reliability for non-ASD group to be ≥ .80.	Not supported
2b) SRS-2 reliability for non-ASD group to be ≥ .90.	Not supported

Table 20 (Cont'd)

2c) 23Q composite reliability for non-ASD group to be lower than that of the SRS-2 and SCQ.	Not supported
3a) SCQ mean of the ASD group to be higher than that of the typically developing group.	Supported
3b) SRS-2 mean of the ASD group to be higher than that of the typically developing group.	Supported
3c) 23Q mean of the ASD group to be higher than that of the typically developing group.	Supported
4a) SCQ to have a sensitivity and specificity of $\geq$ .80.	Supported
4b) SRS-2 to yield a sensitivity and specificity of $\geq$ .80.	Supported for a cut-score of 70, but not a cut-score of 85
4c) The 23Q ASD-related items to have a lower sensitivity and specificity relative to the SCQ and the SRS-2.	Not supported, but much worse at classification overall
5) No specific hypotheses were set for alternative cut-scores.	All three instruments yielded alternative optimal cut-scores (i.e., 10, 62 and 3 for the SCQ, SRS-2 and 23Q, respectively).
6a) Correlations of the SRS-2 and SCQ to be $\geq$ .50.	Supported
6b) Correlations of the SRS-2 and 23Q ASD-related items hypothesized to be lower than the correlations between the SRS-2 and the SCQ.	Not supported when using 10-item 23Q ASD summative scale items (i.e., 14 – 23), because correlation was not significant.
6c) Correlations between the SCQ and the 23Q hypothesized to be lower than the correlation between the SRS-2 and the SCQ.	Not supported when using the 10 item 23Q ASD summative scale (items 14 – 23), because correlation was not significant.

CHAPTER 5: DISCUSSION

The primary purpose of this study was to investigate the psychometric properties of caregiver-rated SCQ (Rutter, Bailey & Lord, 2003), SRS-2 (Constantino & Gruber, 2012), and 23Q (Kakooza-Mwesige et al., 2014) in screening 4-to-18-year-old children for ASD in Uganda. In this chapter, psychometric findings from the SCQ, SRS-2, and 23Q in the present study are compared to results from similar validation studies and put into context within the literature. Study strengths and limitations, clinical implications, and research implications and future directions will also be discussed.

As revealed in the prior chapter, internal consistency reliability findings in the present study sample varied considerably for each of the three instruments, depending on whether the estimate came from the ASD group, non-ASD group, or combined sample. Of the three instruments, the SRS-2 yielded consistently higher reliability across the three conditions. In terms of mean differences between ASD and non-ASD groups, the SCQ and SRS-2 yielded effect size *d* estimates that were > 4.00 in the raw score metric, while the set of 10 ASD items from the 23Q yielded a more modest $d = .58$ effect size. Sensitivity and specificity findings for both the SCQ and SRS-2 were very good (all > .90 for both sensitivity and specificity) based on recommended cut-scores of 15 (SCQ) and 70 (SRS-2 raw score), as well for more optimal sample-specific cut-scores of 10 (SCQ) and 62 (SRS-2 raw score). However, sensitivity and specificity results for the set of 10 ASD items from the 23Q ranged from very poor (for the recommended cut-score of one or more) to modest (.77 sensitivity and .68 specificity based on an optimized sample-specific cut-core of three or more). Finally, the correlation between the SCQ and SRS-2 total scores was high (> .90), while the correlations between the summative

score for the 10 23Q ASD items and both the SCQ and the SRS-2 total scores were low and non-significant (< .20).

Reliability

The SCQ

An SCQ internal consistency reliability of .61 was found in the ASD sample and an estimate of .72 was obtained among those without ASD. For research purposes, a reliability of .70 or higher is considered adequate, while .90 and above is recommended for assessments from which critical clinical decisions are to be made (Nunnally & Bernstein, 1994). Internal consistency estimates of .80 or higher are considered appropriate for screening purposes (Salvia, Ysseldyke, & Bolt, 2010). Thus, the SCQ reliability reported in this study for the ASD sample was below that found by Lord and colleagues (2003), that is, .81 and .86 (autism-specific and broader ASD groups respectively), and .92 among the non-spectrum disorders. This is also below that reported by Bolte and others (2008) for the German SCQ version (i.e., .83 among those with ASD), although they did not report the reliability in the non-autism group.

All other SCQ validation studies reviewed for this purpose reported overall internal consistency reliability based on data from both ASD and non-ASD cases pooled together as one sample--and not separate estimates for ASD and non-ASD. Therefore, further comparisons were made based on this type of overall internal consistency estimate. In the present study, an overall SCQ internal consistency of .93 resulted when ASD and non-ASD cases were combined and treated as one sample. This estimate was similar to those found in studies with the Arabic (.92) and Greek (.91) versions of the SCQ (Aldosari et al., 2020; Zarokanellou et al., 2017), but generally higher than values reported for the Turkish (.89) and German (.83) versions of the SCQ (Avcil et al., 2014; Bolte et al., 2008).

In summary, the internal consistency reliability for the SCQ for children with ASD in the Ugandan context (i.e., based on ASD cases only) was lower than expected and fell short of standards for both research and clinical use. In the non-ASD sample from Uganda, the internal consistency estimate met research standards, but fell short of standards for screening and clinical use. However, when the internal consistency reliability was estimated based on the pooled ASD and non-ASD cases, the SCQ internal consistency estimate was adequate; met standards for research, screening, and clinical use; and was generally comparable to SCQ estimates from other cultural studies.

The SRS-2

The SRS-2 internal consistency reliability estimates among the ASD and non-ASD groups were .84 and .89 respectively, which is adequate for both screening (Salvia, Ysseldyke, & Bolt, 2010) and research purposes--but below the .90 minimum recommended for use in clinical diagnosis or other important decisions about an individual (Nunnally & Bernstein, 1994). It is noteworthy that the internal consistency estimate for the combined sample [i.e., both ASD and non-ASD cases pooled together] was .97, which meets standards for research, screening, and use in individual decision making—though, in practice, one should not use a single measure or source for important clinical decisions. The group-specific internal consistency reliability estimates were also lower than those reported for a larger clinical sample in the SRS-2 manual (i.e., .95 for those with ASD and .97 for unaffected siblings, based on parent ratings; Constantino & Gruber., 2012). However, they are similar to estimates found for the Mandarin version in Taiwan (i.e., .87 and .85 for ASD and control groups, respectively) as well as .89 for non-ASD groups in both the Vietnam and China (Mandarin) studies (Cen et al., 2017; Nguyen et al., 2019; Wang et al., 2012). Also, the reliability found among those with ASD for this study (.84), was

lower than the .96 found in the Vietnamese and German SRS studies, and the .92 result in both the Mandarin and Spanish SRS studies in China and Mexico. The internal consistency among the non-ASD group in the present study (.89) was similar to, though slightly lower than, estimates reported for the German and Spanish SRS versions (i.e., .91 and .92, respectively; Bolte et al., 2011; Fombonne et al., 2012). Though differences in variability play a role in these differences across studies, it is not clear if these differences in sample variability and reliability estimates truly reflect culture-specific or language-specific issues. Yet, these studies provide what are currently our best estimates of SRS reliability in these contexts.

In summary, the SRS-2 internal consistency estimates were lower than expected when calculated separately for ASD and non-ASD groups in the Ugandan context—though standards were met for research and screening use in both groups, and the .97 internal consistency value for the two groups combined met standards for research, screening, and individual decision making. However, the group-specific estimates were consistent with a number of other estimates reported in other cultural contexts, and these were consistently higher than the group-specific estimates obtained for the SCQ in the present study.

The 23Q

This was the first study to examine the internal consistency reliability of a summed subscale made up of the 10 ASD items from the 23Q—in this case, screening children with a known diagnostic status (i.e., either ASD or non-ASD). Internal consistency estimates of .60 and .83 were found for the 10 ASD items of the 23Q among the ASD and non-ASD groups, respectively. When completed by all parents regardless of their child's age, the internal consistency for this 10-item scale (23Q items 14 -23) when evaluated with all ASD and non-ASD cases combined (i.e., with ASD and non-ASD groups combined) was .65. However, the

estimate increased to .92 when only items 19 – 23 (five items more typically administered only to parents of children less than five years of age) were included. In addition, a summed composite consisting of just items 19 – 23 with an optimal cut-score identified more children at risk of ASD than when all 10 ASD items available on the 23Q (i.e., items 14 – 23) were used. It is of critical importance to mention that items like these administered to children less than five years of age on the 23Q (i.e., 19 – 23) are regularly utilized in other instruments as part of the assessment of older children (e.g., SRS-2). Taken together, these findings suggest that, in the context of 23Q usage, items 19-23 on the 23Q should be considered as standard for rating all children regardless of age--and not used exclusively to rate children younger than five years old.

Overall, all three instruments met internal consistency standards for research, screening, and clinical use under at least some conditions (e.g., SCQ in the combined sample, SRS-2 in the combined sample, and 23Q in the combined sample when a subgroup of five items were used). However, the SRS-2 performed relatively better than the other two instruments under all conditions (i.e., ASD only, non-ASD only, and combined samples). Critically, the 23Q 10-ASD-item composite performed poorly overall in terms of reliability—with estimates below .70 in both the ASD-only and combined (ASD and non-ASD combined) conditions. Nevertheless, its performance improved considerably (> .90) in the combined context when a subgroup of five more internal consistent items (23Q items 19-23) was used without the other five items.

Validity

Mean Differences

For all three instruments, and consistent with the ASD construct assessed by each measure, the ASD group scored significantly higher in comparison to the non-ASD group--providing initial evidence of differentiation (clinical discrimination) between clinical and non-

clinical samples. Both the SRS-2 and SCQ total raw scores yielded very large standardized mean differences (i.e., $d > 4.00$ for both instruments) between the ASD and non-ASD group means. The difference between groups remained very large ($d = 3.20$) for the SRS-2 even when norm-referenced T-scores (based on U.S. norms) were used. However, the SRS-2 effect size *d* in the raw score metric ($d = 4.69$) was relatively, and considerably, larger than in the T-score metric ($d = 3.21$)--a difference of almost 1.5 standard deviations. This difference due to metric should be considered when deciding which type of score to use.

Consistent with the SCQ mean difference found in the present study, in the Ugandan context, large mean differences were reported in the SCQ manual, as well as in Greek, Chinese, and Arabic validation studies (Aldosari et al., 2019; Gau et al., 2011; Lord et al., 2003; Zarokanellou et al., 2017). Also, generally consistent with the present study results for the SRS-2, the large SRS-2 mean differences between ASD and non-ASD groups were reported in the SRS-2 manual (i.e., means of 106.6 and 24.6, for ASD and unaffected/control group, respectively; $d = 2.70$; Constantino and Gruber., 2012)--though the exact effect size *d* value was relatively smaller in the SRS-2 manual. Further, large mean differences were reported for SRS-2 studies in Mexico, Germany, and China (Cen et al., 2017; Cholemkery et al., 2014; Fombonne et al., 2012).

For the 23Q, the standardized mean difference between ASD and non-ASD groups in the present study was in the medium range ($d = .58$) for the sum of the 10 ASD items. Though significant, this *d* value was considerably smaller than those observed for the SCQ and SRS-2 and suggests relatively less differentiation between ASD and non-ASD groups for the 23Q items in the Ugandan context. However, it is noteworthy, that when only five items (23Q items 19-23) were used as a scale, the effect size *d* increased to 1.05. This value is still smaller than those for

the SCQ and SRS-2, but it is a considerable improvement over the 10-item version. As noted previously, these five 23Q items are typically administered only when children are under five years old. However, they were completed for children of all ages (both younger and older) in the present study and appeared to function very well within this sample. Thus, consideration should be given to administering these five 23Q items regardless of the child's age.

In summary, all three instruments yielded significant mean differences between ASD and non-ASD groups. However, standardized mean differences were considerably larger for the SCQ and SRS-2 than the 23Q. It was noteworthy that the 23Q's differentiation between ASD and non-ASD improved significantly when only five of the ASD-related items (i.e., items 19-23) were used to construct the scale. Nevertheless, even in this context, the SCQ and SRS-2 between group differences were considerably larger.

Sensitivity and Specificity

The SCQ. The SCQ cut-score of 15 (recommended in the SCQ manual) had excellent sensitivity and specificity (i.e., .96 and .95, respectively) in the present Ugandan sample. These values are higher than those reported for the same cut-score in the SCQ manual (i.e., sensitivity of .85 and specificity of .75; Lord et al., 2003). In terms of the SCQ in other cultural contexts, the Greek version of the SCQ had similarly high sensitivity and specificity estimates (i.e., .96 and .99, respectively). However, this was not the case in the Arabic, Turkish, Chinese, and German SCQ validation studies. The sensitivity and specificity were .80 and .97 for the Arabic version, .71 and .77 for the Chinese version, 1.0 and .33 for the Turkish version, and .89 and .91 for the German version (Aldosari et al., 2019; Avcil et al., 2015; Bolte et al., 2008; Gau et al., 2011; Zarokanellou et al., 2017). Thus, sensitivity and specificity values vary considerably among SCQ validation studies across various cultural contexts. However, the values reported for

the Ugandan (present study) and Greek validation studies stand out high relative to those reported from other contexts.

The SRS-2. According to the SRS-2 manual, raw scores are typically used for large scale screening and reported in research studies involving clinical discriminant work, while norm-referenced T-scores are more typically used for communication in clinical and educational settings (see Constantino & Gruber, 2012). The manual indicates that a variety of different possible cut-scores could be used for screening--depending on the screening context, groups being compared, base rate of ASD in that setting, purpose of a research study, etc. However, a cut-score of 70 (raw score) is recommended for screening minimal risk samples and cut-score of 85 (raw score) for clinical samples where higher confidence in positive screens is desirable (Constantino & Gruber, 2012).

In the present study, these cut-scores were applied in both the raw score and norm-referenced T-score metric. In the raw score metric context, sensitivity was .98 and specificity was .96 for a cut-score of 70, and .84 and 1.00 for a cut-score of 85. In the norm-referenced T-score context (U.S. norms), sensitivity was .86 and specificity was 1.00 for a cut-score of 70, and .28 and 1.00, respectively, for a cut-score of 85. Constantino and Gruber (2012) reported sensitivity of .78 and specificity of .94 for a raw cut-score of 70 in the U.S. context. In the present study sample from Uganda, sensitivity, and specificity values for raw cut-scores of 70 and 85, as well as for the norm-reference T-cut-score of 70, exceeded values reported by Constantino and Gruber for a raw score of 70 based on U.S. data.

Constantino and Gruber (2012) did not report sensitivity and specificity estimates for a cut-score of 85, but they did suggest that its use would generally be in a different clinical context where the base rate of ASD is higher and one is looking for a higher degree of confidence in the

positive screening decision (see p. 21)--though the risk false positives would be reduced, the risk of false negatives may be higher in this situation. When compared to other validation studies that utilized a raw cut-score of 85, the sensitivity (.84) and specificity (1.00) in the present study were both higher than values reported in Taiwanese (sensitivity .66, specificity .89) and German (sensitivity .74, specificity .81) validation studies (Bolte et al., 2008; Wang et al., 2012). No other study reported use of a norm-referenced T-score cut-score of 85 and the low sensitivity of .28 for that cut-score in the present study clearly recommends against it use for general screening.

Overall, raw scale cut-scores of 70 and 85, as well as a norm-referenced T-score cut-score of 85, performed well in terms of general screening sensitivity and specificity in the Ugandan context. However, of the recommended cut-scores, the raw scale cut-score of 70 appeared to balance sensitivity (.98) and specificity (.96) best in this screening context—suggesting good generalizability between the U.S. and Uganda for this recommended cut off.

The 23Q. Results for the 10 ASD items from the 23Q were surprising. The recommended cut-score of one performed very poorly, as it effectively screened in or identified all children in the sample as ASD—regardless of actual group membership (i.e., sensitivity = 1.00, specificity = 0.00). Using a cut-score of one was consistent with the use of these 23Q ASD items in practice, as guidelines indicate an endorsement of any one of these items would lead to a positive screen (Kakooza-Mwesige et al., 2014). Clearly, the SCQ and SRS-2 performed considerably better with their respective recommended cut-scores in the present Ugandan sample. Given that this was the first study to use the ASD items from the 23Q as a summation scale, there were no prior studies available for direct comparison. However, Kakooza-Mwesige et al. (2014) reported sensitivities of .52-.57 and specificity of .92 in general screening for ASD

in Uganda—using the criterion of one or more ASD items being endorsed taken as a positive screen. It is not clear why these ASD items appeared to perform better in this prior study. In that study, there was a very large number of non-ASD cases and only a small number of ASD cases, the items were not aggregated, and it is not clear exactly how the positive screens for non-ASD and neurodevelopmental disabilities were dealt with in the sensitivity and specificity calculations. In the next section, concerning alternative optimal cut-scores, the screening effectiveness of other possible cut-scores were explored for the 23Q ASD items--with better results in the present study sample.

Overall, recommended cut-scores for the SCQ (raw score of 15) and SRS-2 (raw score of 70) performed well in the present study's sample from Uganda. However, the cut-score of one for the summation scale consisting of the 10 ASD 23Q items performed very poorly, as it screened in all cases in the sample regardless of actual ASD or non-ASD group status. Other, potentially more optimal, cut-scores were explored in the next section. However, the recommended cut-scores of 15 for the SCQ and 70 for the SRS-2 appear to work well in the Ugandan context—at least for purposes of differentiating ASD from more typically-developing non-ASD cases.

Alternative Optimal Cut-scores

Given that the recommended cut-scores were established in another cultural context, it makes good sense to examine the screening efficiency of other possible cut-scores in the Ugandan context. Thus, ROC analyses were conducted for each of the three instruments. These analyses suggested one or more potentially more optimal cut-score(s) for each instrument in the Ugandan sample and the findings are discussed below.

The SCQ. As indicated above, the SCQ cut-score of 15, recommended in the manual, demonstrated excellent sensitivity and specificity (i.e., .96 and .95, respectively) in the present study sample. However, a review of the SCQ ROC coordinates indicated that, in this Ugandan sample, a revised cut-score of 10 identified all ASD children and had a specificity of .93. Although 10 as a cut-score identified all children with ASD, the specificity dropped slightly--but to an acceptable degree relative to the sensitivity gain (virtually eliminating false negatives for ASD). When compared to other studies in which cut-scores lower than 15 were proposed (e.g., cut-score of 12 in the Chinese and Arabic validation studies [Aldosari et al., 2013; Gau et al., 2011] and 14.5 for a Turkish study [Avcil et al., 2015]), the sensitivity and specificity found among Ugandan children for the alternative cut-score of 10 were higher.

The finding of a slightly improved screening performance by a cut-score that is five points lower than the standard recommendation is very suggestive. However, the relatively small sample size of the present study should be considered and makes these findings more tentative. Because of this, SCQ cut-scores between 10 and 15 should be examined in further Ugandan studies with larger samples.

The SRS-2. For the SRS-2, lowering the cut-score from 70 to 62 (raw score) enabled all children with ASD in the present study sample to be identified (i.e., sensitivity of 1.0) with the same specificity as attained when utilizing both the raw and standardized cut-score of 70 (i.e., .96). It is noteworthy that this cut-score of 62 also performed well in a clinical study (ASD vs. unaffected siblings) reported in the manual for the SRS-2 (sensitivity and specificity of .92; Constantino & Gruber, 2012, p. 50). Furthermore, a raw cut-score of 62 was recommended based on results (sensitivity of .93, specificity of .98) in a Vietnamese study (Nguyen et al., 2019). Compared to both other studies (from the test manual and from Vietnam), the sensitivity of the

cut-score of 62 in the present Ugandan sample was higher. However, specificity in the Ugandan sample (.96) fell between the values reported in the clinical study (.93) from the SRS-2 manual and the study in Vietnam (.98)--suggesting a generally similar range of specificity across the three studies that used this cut-score. These findings suggest that the more optimal cut-score of 62 found in the present Ugandan study sample may have some generalizability across cultural contexts. However, these findings should not be over interpreted, as most SRS-2 validation studies from other countries that were reviewed for this purpose, proposed different cut-scores than those recommended in the SRS-2 manual. Cut-score recommendations ranged from 43 for children in Germany (Cholemkery et al., 2014) to 87 among children in Taiwan (Wang et al., 2012). Only Bolte and colleagues (2008) found a cut-score of 100 for clinical groups in a German validation study. Despite their recommendation, the cut-score of 100 had a low sensitivity of (i.e., .56) when compared to that of the present study (i.e., 1.0) with a cut-score of 62. However, specificities from both studies were in excellent ranges, that is, .90 for the German validation and .96 in the present Ugandan study—though resulting from very different cut-scores.

The 23Q. The 23Q cut-score of 3 or more yielded a sensitivity and specificity of .77 and .68 respectively. When compared to the recommended cut-score of 1 with .96 and 0.0 for sensitivity and specificity respectively, the cut-score of 3 provided a much better balance of these screening indices. (Using a cut-score of 4 would improve specificity to .80 but would reduce sensitivity to .55--meaning 80% of those without ASD would be accurately screened out while 45% [almost half of participants with ASD] would be missed by the screener.) Kakooza-Mwesige et al. (2014) reported sensitivities of .52-.57 and specificity of .92 for the 10 ASD items ($\geq$ 1 item endorsed = positive screen) in a general sample from Uganda—with only a small

number of ASD cases. In the present study, using the cut-score of one was not useful at all for screening, while a cut-score of 3 yielded greater sensitivity and poorer specificity than Kakooza-Mwesige et al. (2014) who counted any of the 10 ASD items being endorsed as a positive screen. It is possibly noteworthy that the age range of participants differed across the two studies (two to nine years old in Kakooza-Mwesige et al., and four to 18 years old in the present study). This age difference may have impacted findings. However, it is not fully satisfying as an explanation, because the cut-score of one performed so poorly in the present study despite the inclusion of many children from within the same age range as Kakooza-Mwesige et al. (e.g., the 0.0 specificity in the present study vs. the .92 specificity in the Kakooza-Mwesige et al. study cannot be explained purely on the basis of the age range difference when the same cut-score of one or more was used). Regardless of why the results were so different across the two studies, in a large screening situation (i.e., screening hundreds or thousands of cases), the result in the Kakooza-Mwesige et al. study suggests that an unacceptable percentage of those with ASD would be missed by the screener, while the results of the present study suggest that the optimum cut-score would likely result in too many false positives in the general population. These ASD items should be considered for revision and re-examined in further studies.

Overall, ROC analyses revealed that cut-scores of 62 on the SRS-2 and 10 on the SCQ may improve their screening effectiveness in the Ugandan context. These cut-scores performed slightly better, in the present Ugandan sample, than the recommended cut-scores of 70 (SRS-2) and 15 (SCQ). However, it is noteworthy that the recommended cut-scores also performed well (at levels close to the more optimal cut-scores) and the new cut-scores derived from the present sample data may require replication in larger and broader samples, to assure stability and generalizability, before being recommended for more general use in Uganda. Results from ROC

analyses of the 10 ASD items from the 23Q suggested that, in the present sample, even best compromise cut-score of 3 (sensitivity .77, specificity .68) would still fail to screen out sufficient cases without ASD to be considered an efficient screener--resulting in too many false positive cases that would require follow-up evaluations. Therefore, the SRS-2 and SCQ, modified for the Ugandan context, appear appropriate for use. However, the ASD items from the 23Q are more problematic and likely require revision to improve sufficiently for regular use.

Correlations Among Study Instruments

Convergent validity among the three instruments was examined via Pearson correlations between their overall composite scores. The positive correlation between the SCQ and SRS-2 was very high (i.e., $r = .93$, $p < 0.001$), while the correlations between the sum of the 10 ASD items from the 23Q and the other two instruments were low and non-significant ($r = .19$, $p = .053$ with SCQ; $r = .17$, $p = .082$ with SRS-2). However, when the correlations were run again using just items 19-23 from the 23Q, the correlations involving the 23Q improved to moderate and significant (i.e., $r = .35$, $p = 0.001$ with the SRS-2 and $r = .36$, $p < .001$ with the SCQ). The convergent validity issues found here for the sum of the 10 ASD items from the 23Q are further evidence of the need to closely re-examine and revise these items.

Bolte and colleagues (2008), in a German validation study, were among the few authors who reported a correlation between the SCQ and SRS total scores (i.e., $r = .58$). These authors reported that this correlation was obtained from a sample of 119 “probands”. Most of the other validation studies reviewed for this purpose administered only one of the instruments (i.e., the SRS or SCQ), in addition to other instruments, but not both. The correlation between the SRS-2 and SCQ for this study was considerably stronger than that reported by Bolte and colleagues, which could be due to sample differences (e.g., inclusion of both ASD and non-ASD cases in the

present sample, proband-only sample in the Bolte et al. study). Additional estimates of this relationship would be helpful in future samples to provide more context and suggest possible moderating conditions that may impact the magnitude of the correlation.

This study is the first to report correlations between 23Q summed ASD item sets and composites from the SRS-2 and SCQ. As indicated above, the low convergent correlations involving the 23Q ASD items suggest the need to re-examine these items with an eye toward revision and further research.

Strengths and Limitations

The present study involved a number of strengths, which highlight its importance and quality of its findings. First and foremost, this is the first study to examine the validity of standardized western instruments, such as the SCQ and SRS-2, for ASD screening in Uganda. The need for high quality screening instruments in Uganda is high, as available local options are limited. In addition, ASD-related items from the 23Q were also evaluated and compared to the SCQ and SRS-2 in the present study. This was important, because this survey screening instrument was previously used in Uganda. However, its performance in the present study was much poorer and more problematic when compared to the SCQ and SRS-2—especially in terms of screening sensitivity and specificity, and convergent validity.

Second, the simultaneous psychometric evaluation of three different screening instruments within the sample was a clear strength. This allowed for assessment of all instruments in terms of internal consistency, discrimination between ASD and non-ASD cases according to both recommended and sample-specific cut-scores, and convergent validity (i.e., correlations among the three instruments). Evaluating them simultaneously in the same sample

allowed for direct comparison of the instruments in the same context, with the same cases, and using the same criteria.

Third, permission from the copyright holder of the SCQ and SRS-2 (Western Psychological Services) was obtained and the test company assisted the author's efforts to identify items with potentially confusing cultural-laden content and, when necessary, to adapt the content to fit the cultural context in order to maximize appropriateness, relevance, and rater comprehension. Items were subject to review by consultants in Uganda (in consultation with the author and feedback from the test company experts), modified for enhanced cultural relevance—with suggestions from the consultants, and piloted with feedback from a group of potential participants. This feedback was used to identify potentially confusing items, consider further modification and/or to craft standardized, brief clarifications that the interviewer could use when requested. This was all done prior to recruitment of actual participants and execution of the study.

Fourth, children across a relatively wide age range were included in the study (i.e., ranging in age from four years to 18 years old). This encompasses the full age range for the SRS-2 school-age form and falls within ages appropriate for the SCQ. The 23Q was previously examined in Uganda with children ages two to nine years old (Kakooza-Mwesige et al., 2014). This study partially overlaps with and expands this age range.

Fifth, trained clinical psychologists with experience working with children who have ASD were recruited to administer the study instruments to parent participants. Such experience and training made the interviewers well suited for developing rapport, providing standard item explanations when needed, and for providing expert review of diagnostic records to evaluate ASD cases for study inclusion.

Sixth, the use of DSM-5 diagnosis of ASD as the criterion for distinguishing between ASD and non-ASD groups was also a strength. “Gold standard” diagnostic instruments, such as the ADI-R and ADOS-2, are not readily available in Uganda. However, the DSM-5 is used consistently in major hospitals and other diagnostic settings within the country—and assuredly within the recruitment area for the study. Records were subject to review by the clinical psychologists employed by the study in Uganda to assure that the DSM-5 was used to diagnose cases—which was required for inclusion in the study’s ASD group. (Cases diagnosed without clear use of DSM-5 criteria were not included in the study—in either group.) This provided a consistent diagnostic standard, which was the best available within the country.

Despite the strengths of this study discussed above, there were also number of limitations. First, the sample size for the study (N = 107) was relatively small when compared to many other validation studies. This can limit the precision of estimates derived from the sample and estimate stability across samples. This is, in part, why the lower-cut-scores derived from the ROC analyses require cross-validation to improve confidence.

Second, lack of greater geographic diversity in the sample from across Uganda is a limitation. All study participants lived within the Kampala or Wakiso districts. Thus, generalization of results may be restricted to these regions.

Third, language diversity is restricted. English and a variety of Ugandan dialects are spoken across Uganda. The majority of people around the major cities speak English and the diversity of Ugandan dialects would limit the applicability of a single Ugandan dialect translation and adaptation of the instruments. Thus, English was chosen as the primary language for the cultural adaptation of the instruments.

Fourth, it was necessary to deviate from the typical expectation of having parents read and complete the rating scale instruments on their own. The relatively high rate of illiteracy worked against this procedure. To adapt, the instructions and items from the rating scales were administered to all parent participants via interview. This kept the procedure uniform, accommodated those whose reading skills were not sufficient to validly complete the scales, and readily allowed for standardized item clarifications from the interviewer, if needed. This procedure was very appropriate for this cultural context. However, the standardized administration does deviate significantly from having participants read the items themselves with the typical expectation of valid comprehension. This is a generalization issue, but one that appears necessary to allow for such instruments to be broadly and meaningfully used within this current cultural context.

Fifth, though DSM-5 diagnostic criteria were applied across the sample to establish the ASD diagnosis, no "gold standard" instrument (e.g., ADI-R or ADOS-2) was used to confirm diagnosis. These instruments are not readily available or used within Uganda. In fact, there is a need to adapt such instruments for use in Uganda, as no validated adaptations of these instruments currently exist in there.

Sixth, the clinical discriminant validity of these instruments was examined in the present study only for differentiating between ASD and generally typically-developing non-ASD groups. The differentiation between these two groups is a good start for this type of clinical discriminant validity and is generally applicable for screening purposes. However, it is very important to understand how well these instruments discriminate between ASD and clinical groups, and especially between ASD and other types of non-ASD developmental disabilities. These types of comparisons more closely match the types of differentiations that clinicians are regularly asked

to make in the context of diagnosis. Other clinical groups or other types of developmental disability groups were not included in the present study.

Clinical Implications

The current study is a pioneering investigation of screening instruments that could be used to detect symptoms of ASD in Ugandan children. Results of this study support the contention that both the SCQ and SRS-2 are potentially viable instruments needed to screen for ASD in the Ugandan context. However, further research studies are needed to confirm and expand upon these psychometric findings before full confidence in the use of these instruments is warranted.

With these preliminary data, the combined sample of ASD and non-ASD showed reasonable reliability for both the SCQ and the SRS-2. However, the SRS-2 was relatively more reliable when compared to the other instruments when used separately on ASD and non-ASD. Thus, professionals must be mindful of how the reliability may be affected by the population of children to which an instrument is applied. In addition, although the 10 ASD items of the 23Q yielded low reliability as a scale, the five items originally intended only for younger children yielded considerably higher internal consistency reliability as a scale when applied across a broader range of ages. This suggests the possibility of improving the instrument for applied screening by including more developmental history items. However, the current 23Q ASD items are not recommended for use in clinical contexts.

When used in differentiating ASD from non-ASD samples among the Ugandan children in this study, the SCQ and the SRS-2 showed considerably larger effect sizes compared to the 23Q, indicating that both the SRS-2 and the SCQ have stronger discriminant validity for differentiating between cases with ASD and typically-developing non-ASD cases—a distinction

similar to broad screening in the general population. Results of the ROC analyses showed that a lower cut-score of 62 enhanced the sensitivity and specificity of the SRS-2 in this sample, but that the recommended screening cut-score of 70 was also acceptable. In addition, a lower cut-score of 10 of the SCQ performed better in this study--although the recommended cut-score of 15 was also acceptable. In general, the two recommended cut-scores appear to be safer to apply in practice until the lower cut-scores are either supported or refuted by replication studies and studies involving other types of groups. It is very likely that a different cut-score may be more useful for different purposes (i.e., different differentiation needs). Further research in the Ugandan context is needed to clarify these possible situational variations.

Research Implications and Future Research Directions

Up until now, no formal validation studies have been done to examine the psychometric properties of established instruments from the U.S. and Europe to screen for children for ASD symptoms in Uganda. Thus, findings from this study provide a foundation to explore viable instruments, i.e., the SRS-2 and the SCQ, that are well validated in the U.S. and that may be used in the Ugandan culture with some modifications. The screening utility of the 23Q, an instrument that was originally developed for the African culture to screen in the general population for children with developmental disabilities, including 10 items related to ASD, was further explored. As mentioned above, instruments imported from another culture, like the SRS-2 and SCQ, may require modifications of items and/or utilization in order to reliably and accurately screen children with ASD in the Ugandan culture. However, as addressed in the limitations section, a combination of larger sample size, broadening the cross-group comparisons to include other clinical groups, and perhaps expanding investigations of important culture-specific

elements can provide further information to improve these screening instruments for Ugandan children with ASD.

Though this study showed that the SRS-2 and the SCQ are potentially viable instruments for use in the Ugandan context, it is the first study to be conducted. Cross-validation of ROC findings to better understand the generality of revised cut-scores in the Ugandan context is particularly important and this will be further enhanced by the incorporation of other relevant clinical groups in these studies for comparisons that reflect the kinds of differential diagnostic distinctions that clinicians most frequently face. Such differential diagnoses may involve such distinctions as ASD vs ADHD, intellectual disability, speech/language disorders, anxiety disorders, etc. The robustness of an instrument for making such distinctions is important to understand and can suggest avenues for instrument improvement across successive revisions. Furthermore, the 10 ASD items from the 23Q provided a foundation for ASD screening in Africa, the present study suggested significant weaknesses in the present item set that imply a need for revision and further evaluation of items to improve the reliability and screening effectiveness of the instrument.

The issue of sample size is important for obtaining more stable estimates involving less sampling error. All other things being equal, larger samples in psychometric studies instill more confidence in the stability of the findings across similar samples and contexts. Thus, future validation studies for the SRS-2 and SCQ in Uganda should strive for significantly larger samples for cross-validation purposes.

Another important issue is the broadening of sample representation to reflect a broader range of population characteristics (e.g., geographic regions, socioeconomic range, other settings [e.g., non-hospital], etc.) in order to better understand the generality of the findings. In addition,

better sample representation can improve the accuracy of population estimates. Ultimately, larger and more representative sample reflect appropriate movement toward the longer-term goal of instrument norming. It will ultimately be helpful to establish Ugandan norms for instruments like the SRS-2 in order to allow for standardized reference scores to be used to communicate results in clinical and educational settings. However, this type of activity should follow the proper establishment of the relevance and validity of the instrument in Uganda for screening, clinical, and educational purposes.

Though the sample in the present study is generally well characterized demographically, to the extent it is possible, obtaining cognitive and language scores, as well as implementing a stronger autism diagnostic approach would enhance the quality of studies in this context. Potentially viable cognitive and language instruments exist with some validity for this context, but finding options approaching a "gold standard" diagnostic instrument (e.g., ADI-R or ADOS-2) may not be reasonable until similar studies are performed to establish the generalization of ADI-R and ADOS-2 diagnostic validity in the Ugandan context. In other words, establishing the validity of rating scale instruments, such as the SRS-2 and SCQ, for ASD screening, may be first steps toward validating a more comprehensive ASD diagnostic instrument for use in Uganda.

Further explorations of the correlations across different ASD screening instruments, between screening instrument results and DSM-5 diagnosis, and between screening instruments and established diagnostic instruments (e.g., ADI-R and ADOS-2, which may also require validation work in Uganda before their wider use becomes an option) are important for understanding the larger scope of construct validity for an instrument or the multiple instruments and techniques involved. In addition, it may be fruitful to explore how multiple screening instruments might be used more effectively together (e.g., compensate for each other's screening

strengths and weaknesses, assess differential weighting options with the goal of improving the accuracy of screening decisions beyond what is possible for a single measure, etc.).

The SRS-2 and SCQ were designed to be completed by a primary caregiver or teacher familiar with the child being assessed. In this study, a trained researcher read the items to caregivers of the child as opposed to self-administration by the caregiver or teacher. Therefore, the potential differences in the administration method should be further examined. This may be possible only within samples of literate adults, which may not generalize as well to adults with lower literacy skills, as such skills may be correlated with other important variables. However, the findings from literate samples wherein the different approaches to administration are compared can be very informative.

Finally, only a primary caregiver of a child with ASD completed the instruments in the current study. Having multiple informants across different settings that can rate the children, e.g., teachers at school can improve the generalizability of the results across different settings. In addition, having multiple raters who know the child well rate the child high on an instrument like the SRS-2 statistically increases the likelihood of a confirmed ASD diagnosis later (Constantino & Gruber, 2012). Thus, validity data for multiple rater types can contribute meaningfully to the screening utility of the measure.

Conclusions

Results of the present study lend initial support for the viability of SRS-2 and SCQ versions culturally adapted for use in Uganda. These preliminary results support the screening effectiveness of the SRS-2 and the SCQ for distinguishing between cases with ASD and those more typically-developing non-ASD cases. This situation is similar to that which occurs in broader general population screening. However, both replication in larger, more representative

samples and discriminative results for ASD vs other clinical groups are essential for understanding the robustness of these initial findings and the broader applicability of these instruments. Based on the results for the 10 ASD items from the 23Q, this measure is not recommended for critical screening situations. Though the instrument was initially designed for broad screening use in Uganda, the sensitivity and specificity results, in particular, from the present study indicate that its recommended cut-score performed very poorly. Furthermore, though a more optimal cut-score found through ROC analyses indicated a composite consisting of these 10 23Q ASD items can accurately screen in approximately 3/4s of those cases with ASD, even when using this improved cut-score it will also miss-classify almost 1/3 of non-ASD cases as ASD. The screening effectiveness of the SRS-2 and SCQ was considerably better within the same sample. It is recommended that the 23Q ASD items be re-examined and revised going forward to improve its screening effectiveness.

APPENDICES

APPENDIX A: Table 21

Table 21. *Reliability of SCQ Total Scores Among Non-ASD Group with "Compulsions and Rituals" (Item, 19) Deleted*

	Item	Item-Total Statistics			
		Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
2	To and fro	5.29	13.371	.303	.720
3	Said same	5.02	13.036	.170	.726
4	Inappropriate	4.91	13.174	.112	.732
5	She/he I	5.23	13.491	.118	.726
6	Metaphorical	5.16	13.519	.061	.731
7	Same exactly	5.13	12.948	.247	.720
8	Ritually	5.18	13.495	.078	.729
9	Facial expressions	5.2	13.397	.128	.726
10	Hand a tool	5.05	13.288	.103	.731
11	Preoccupied	5.16	13.483	.074	.730
12	Interest parts	5.21	12.862	.385	.713
13	Interests age	5.2	12.961	.312	.716
14	Sight, feel, sound,	5.23	13.636	.049	.729
15	Hand and fingers	5.25	13.427	.172	.724
16	Body movements	5.29	13.626	.116	.726
17	Attachment	5.25	13.755	.000	.731
18	Talking friendly	5.18	12.804	.354	.713
20	Spontaneous copying	5.23	13.6	.066	.729
21	Spontaneous point	4.89	12.679	.251	.720
22	Point/pull	4.73	12.345	.353	.712
23	Head yes	5.18	12.44	.505	.704
24	Shake head	5.07	12.431	.389	.709
25	Directly face	5.3	13.815	-.011	.728
26	Show	5.25	13.355	.211	.722
27	Share other than	5.16	13.446	.088	.729
28	You to join	5.3	13.706	.099	.726
29	To comfort	5.18	13.058	.250	.719
30	Gesture	5.05	12.524	.347	.712
31	Range expressions	5.25	13.245	.269	.719
32	Spontaneously	5.29	13.081	.522	.713
33	Make-believe	5.21	12.644	.487	.707
34	Interest children	5.27	13.363	.245	.721

Table 21 (Cont'd)

35	Responds positively	5.29	13.444	.250	.721
36	Attends to voice	5.2	13.033	.281	.718
37	Imaginatively	5.2	12.561	.486	.706
38	Cooperatively	5.27	12.963	.493	.711

Note. Item stems are truncated to respect the publisher's copyright.

APPENDIX B: Table 22

Table 22. *SRS-2 Total Scores Reliability Among Non-ASD Group if Item is Deleted*

Item-Total Statistics					
		Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
1	Fidget	25.85	243.015	.03	.891
2	Facial match	26	240.667	.12	.890
3	Self-confident	26	236.778	.474	.887
4	Rigid stress	25.78	238.174	.215	.889
5	Doesn't tricks	25.44	242.621	.016	.893
6	Alone	26.02	242.722	.051	.890
7	Others think	25.04	236.517	.209	.890
8	Strange	26.16	241.176	.32	.889
9	Clings	25.84	232.88	.483	.886
10	Literal	25.93	232.958	.542	.886
11	Confidence	25.64	233.384	.386	.887
12	Feelings	25.62	231.833	.424	.887
13	Awkward turn-taking	25.89	230.543	.492	.886
14	Coordinated	26.04	233.184	.608	.885
15	Tone voice	25.84	234.176	.461	.886
16	Eye contact	26.05	237.978	.385	.888
17	Unfairness	25.55	230.141	.513	.885

Table 22 (Cont'd)

18	Difficulty friends	25.96	231.332	.610	.885
19	Gets frustrated	25.93	230.884	.595	.885
20	Unusual interests	26.11	238.877	.324	.888
21	Imitate others	25.6	231.985	.409	.887
22	Plays appropriately	25.96	233.406	.481	.886
23	Does not join	25.76	245.11	-.082	.892
24	Difficult routine	26.07	237.809	.409	.888
25	Out step	26.11	241.284	.201	.889
26	Comfort sad	25.53	229.772	.531	.885
27	Avoid social	25.8	235.385	.379	.887
28	Same over and over	26	237.296	.374	.888
29	Odd	26.07	235.328	.539	.886
30	Busy place	26.07	240.772	.254	.889
31	Mind off	25.78	241.137	.134	.890
32	Hygiene	25.64	237.051	.288	.888
33	Socially awkward	26.18	240.559	.456	.888
34	Avoids close	25.95	244.423	-0.053	.891
35	Following conversation	25.98	230.092	0.652	.884
36	Difficult adults	25.85	239.83	0.2	.889
37	Difficulty relating	25.91	227.899	0.708	.883
38	Appropriate mood	25.49	230.44	0.469	.886

Table 22 (Cont'd)

39	Limited range of interests	26.04	236.554	0.436	.887
40	Is imaginative	25	227	0.496	.885
41	Wanders aim	25.76	235.369	0.348	.888
42	Sounds	25.95	242.978	0.028	.891
43	Separates easily	25.2	235.163	0.193	.892
44	Understand events	25.64	232.828	0.365	.888
45	Attention away	25.69	239.81	0.176	.890
46	Serious expressions	25.64	243.088	0.003	.892
47	Too silly	26.16	237.917	0.581	.887
48	Jokes	25.42	231.803	0.388	.887
49	Well at few	25.45	236.29	0.252	.889
50	Flap	26.22	242.211	0.386	.889
51	Around subject	25.96	234.962	0.446	.887
52	Loud or noisy	25.4	228.393	0.461	.886
53	Tone voice	26.15	238.978	0.342	.888
54	People objects	26.18	241.263	0.356	.889
55	Knows close	25.09	229.677	0.414	.887
56	Between people	25.78	242.618	0.03	.891
57	Gets teased a lot	25.98	235.685	0.436	.887
58	Much on parts	26.02	239.981	0.217	.889
59	Suspicious	25.98	239.87	0.232	.889

Table 22 (Cont'd)

60	Emotionally distant	26.00	244.185	-0.037	.891
61	Inflexible	25.82	235.04	0.38	.887
62	Illogical for doing	25.91	239.899	0.169	.890
63	Others unusually	26.15	241.423	0.26	.889
64	Relax	26	239.148	0.245	.889
65	Space	26.11	241.284	0.235	.889

Note. Item stems are truncated to respect the publisher's copyright.

APPENDIX C: IRB Clearance in Uganda

MAKERERE UNIVERSITY
P.O. Box 7072 Kampala, Uganda
E-mail: healthsciences.irb@gmail.com
deanshs@chs.mak.ac.ug
Tel: 256 0200903786
Tel: 256 414 531533
Fax: 256 414 531533

COLLEGE OF HEALTH SCIENCES
SCHOOL OF HEALTH SCIENCES
OFFICE OF THE DEAN

June 26th, 2019

Category of review
[X] Initial review
[ ] Continuing review
[ ] Amendment
[ ] Termination of study
[ ] SAEs

Mr. Jorem Awadu Emillian
Department of Counseling Education Psychology
and Special Education
College of Education
Michigan State University, USA

Dear Mr. Awadu,

Re: Approval of research protocol #SHSREC REF: 2019-041
"Validation of Autism Assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM-5 in Assessing for Autism Spectrum Disorder (ASD) in Uganda"

Thank you for submitting an application for ethical review of the above referenced research protocol. The committee reviewed it and granted approval for **one (1) year**, effective **June 26th, 2019.** Approval is valid until **June 25rd, 2020.**

Continuing Review
In order to continue working on this study (including data analysis) beyond the expiration date, the School of Health Sciences Research and Ethics Committee must reapprove the protocol after conducting a substantive, meaningful, continuing review.
This means that you must submit a continuing report form as a request for continuing review. To best avoid a lapse, you should submit the request six (6) to eight (8) weeks before the lapse date. Please use the forms supplied by our office.

Amendment Review
During the approval period, if you propose any change to the protocol such as its funding source, recruiting materials, or consent documents; you must seek School of Health Sciences Research and Ethics Committee approval before implementing it.
Please summarize the proposed change and the rationale for it in a letter to the School of Health Sciences Research and Ethics Committee. In addition, submit two (2) copies of an updated version of your original protocol application- one showing all proposed changes in bold or 'track changes,' and the other without bold or track changes.

Reporting
Other events which must be reported promptly in writing to the School of Health Sciences Research and Ethics Committee include:
Suspension or termination of the protocol by you or the grantor

Unexpected problems involving risk to participants or others

Adverse events, including unanticipated or anticipated but severe physical harm to participants.

Monitoring audit of research study activities

As per the Uganda National Guidelines for Research Involving Humans as Research Participants, Section 3.5, The Research and Ethics Committee has a duty to ensure that all research studies it approves are conducted in accordance with the research governance code of practice. In order to ensure compliance with scientific and ethical requirements, the School of Health Sciences Research and Ethics Committee undertakes random monitoring audits. If your research study is selected for monitoring audit, you will be given three (3) week's notice to prepare all documentation for inspection, Therefore, expect the monitoring team at your study site anytime.

It is your responsibility to inform us in the event of early termination of the research project or if you fail to complete the research project.

Documents approved for use along with the research protocol include:

- Verbal assent script (8-12 years) (English version)
- Translated verbal assent script (8-12 years) (Luganda version)
- Verbal assent script (5-7 years) (English version)
- Translated verbal assent script (5-7 years) (Luganda version)
- Parental informed consent form (English version)
- Social communication questionnaire-SCQ (English version)

Note: Only stamped verbal assent script forms, parental informed consent form and data collection form should be used for data collection. Any data collected using unstamped forms will be considered invalid.

Do not hesitate to contact us if you have any questions. Thank you for your cooperation and commitment to the protection of human subjects in research.

Final approval is to be granted by Uganda National Council for Science and Technology.

Yours sincerely,

MAKERERE UNIVERSITY
SCHOOL OF HEALTH SCIENCES
APPROVED
VALID UNTIL
25 JUN 2020
RESEARCH & ETHICS COMMITTEE
P. O. BOX 7072, KAMPALA

Dr. Paul Kutyabami
Chairperson, School of Health Sciences Research and Ethics Committee
College of Health Sciences, Makerere University

APPENDIX D: MSU IRB Clearance

MICHIGAN STATE
UNIVERSITY

Continuing Review APPROVAL
Pre-2018 Common Rule

February 26, 2019

To: Ka Lai Gloria Lee

Re: **MSU Study ID:** STUDY00000107
IRB: Social Science / Behavioral / Education Institutional Review Board
Principal Investigator: Ka Lai Gloria Lee
Category: Expedited 7
Submission: Continuing Review CR00000578
Submission Approval Date: 2/26/2019
Effective Date: 2/26/2019
Study Expiration Date: 2/25/2020

Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda

This submission has been approved by the Michigan State University (MSU) Social Science / Behavioral / Education Institutional Review Board. The submission was reviewed by the Institutional Review Board (IRB) through the Non-Committee Review procedure. The IRB has found that this study protects the rights and welfare of human subjects and meets the requirements of MSU's Federal Wide Assurance (FWA00004556) and the federal regulations for the protection of human subjects in research (e.g., pre-2018 45 CFR 46, 28 CFR 46, 21 CFR 50, 56, other applicable regulations).

How to Access Final Documents
To access the study's final materials, including those approved by the IRB such as consent forms, recruitment materials, and the approved protocol, if applicable, please log into the Click™ Research Compliance System, open the study's workspace, and view the "Documents" tab. To obtain consent form(s) stamped with the IRB watermark, select the "Final" PDF version of your consent form(s) as applicable in the "Documents" tab. Please note that the consent form(s) stamped with the IRB watermark must typically be used.

Continuing Review: IRB approval is valid until the expiration date listed above. If the research continues to involve human subjects, you must submit a Continuing Review request at least one month before expiration.

Modifications: Any proposed change or modification with certain limited exceptions discussed below must be reviewed and approved by the IRB prior to implementation of the change. Please submit a Modification request to have the changes reviewed. If changes are made at the time of continuing review, please submit a Modification and Continuing Review request.

Office of Regulatory Affairs
Human Research Protection Program

4000 Collins Road
Suite 136
Lansing, MI 48910

517-355-2180
Fax: 517-432-4503
Email: irb@msu.edu
www.hrpp.msu.edu

APPENDIX E: Permission to Conduct Research at Special Children's Trust, Uganda

Special Children's Trust

P.O. BOX 9295; Kampala
Email: sctuganda@gmail.com
dawnspecialchildren@gmail.com

Mobile: 0705445565 / 0393108942
www.specialchildrenstrust.org

DAWN
FUN LEARN and FLOURISH TOGETHER

Reaching out in love, offering hope and restoring dignity to children with special needs in Uganda.

Monday 22nd July 2019

Dear Jorem Emmillian Awadu,

RE: REQUEST FOR RESEARCH ACCEPTED

Following our discussions we would like to give you permission to conduct your research with Special Children's Trust.

However, we have the following conditions;

- Special Children's Trust will be named as a source of your research findings within your final thesis.
- We will receive a copy once it has been approved.
- Children and parents names will not be disclosed and no pictures will be taken.

If you are happy to proceed with these conditions please sign below. We would be happy to work with you.

Yours sincerely

SPECIAL CHILDREN'S TRUST
22 JUL 2019
P. O. BOX 9295 KAMPALA (U)

Sarah Ampumuza
International Coordinator

Jorem Emmillian Awadu

APPENDIX F: Permission to Conduct Research with Children at Butabika National Hospital, Uganda

TELEPHONE: DIRECT 041 - 504376
GENERAL 041 - 504338
FAX NO. 256 -41-504760
E-MAIL: buthosp@infocom.co.ug

THE REPUBLIC OF UGANDA

BUTABIKA HOSPITAL
P. O. BOX 7017
KAMPALA, UGANDA.

IN ANY CORRESPONDENCE ON

THIS SUBJECT PLEASE QUOTE .....

July 16, 2019

Mr. Jorem E. Awadu
Principal Investigator
Department of Counseling, Education Psychology
& Special Education
Michigan State University
East Lansing, MI, 48824

RE: INSTITUTIONAL APPROVAL FOR RESEARCH

We have received your request for permission to carry a research study entitled, **"Validation of Autism Assessment: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM-5 in Assessing for Autism Spectrum Disorder (ASD) in Uganda"** a case study Butabika hospital. We have noted that the research protocol has been approved by Makerere University, School of Health Sciences, Research and Ethical Committee **(SHSREC REF: 2019-041)**

You are granted institutional permission to undertake the research for the period a period of one year after final approval and permission has been granted by the Uganda National Council of Science and Technology (UNCST). We request you to abide by the regulations as stipulated by the research ethical committee and institutional guidelines.

We also request that a copy of your research findings be disseminated to Butabika hospital, which institution should also be acknowledged in all publications.

Yours Sincerely,

MINISTRY OF HEALTH
P. O. BOX 7017
KAMPALA
Date: 16/7/19
BUTABIKA HOSPITAL

Dr. H. Birabwa-Oketcho
HEAD TRAINING/ BUTABIKA HOSPITAL

APPENDIX G: Fidelity Checklist

Fidelity Checklist (Training and Field Testing)

Ratings: 0=Did not implement; NA; 1=Partially implemented; 2=Implemented

Items	Ratings
The administrator greets the caregiver.	0 1 2 NA
The administrator situates the caregiver in a comfortable and private space for the assessment.	0 1 2 NA
The administrator explains the instructions clearly to the caregiver.	0 1 2 NA
The administrator allows time for the caregiver to ask any questions before starting the assessment.	0 1 2 NA
The administrator reads all items to the caregiver clearly.	0 1 2 NA
The administrator is able to read the items at a pace that is appropriate for the caregiver.	0 1 2 NA
If the caregiver has a question or inquiry of any items during the assessment, the administrator is able to explain further to the caregiver without deviating from the protocol to affect the responses.	0 1 2 NA
If the administrator needs to provide further inquiry, he/she documents the necessary information in space provided in the protocol, i.e., what the inquiry is, what the explanation is given, how the inquiry is resolved (e.g., language use, switch to local language for explanation of certain terms).	0 1 2 NA
Upon completion of all the assessment, the administrator takes the time to check all the items in each survey to ensure they are all completed.	0 1 2 NA
If any of the items are not answered, the administrator clarifies with the caregiver.	0 1 2 NA
Before departure, the administrator provides the incentive to the caregiver.	0 1 2 NA
The administrator walks the caregiver out to the testing space.	0 1 2 NA
If the caregiver discloses any information that poses danger, the administrator is able to take immediate actions.	0 1 2 NA
If the caregiver discloses the need for further information (e.g., referrals, resources), the administrator is able to take the necessary actions to gather information and follow up with the caregiver promptly.	0 1 2 NA
Overall, the administrator is able to build positive rapport with the caregiver.	0 1 2 NA
Overall, The administrator interacts with the caregiver in a professional manner.	0 1 2 NA

Note:
For training purposes, a fidelity of at least 95% must be achieved.
For field testing, a fidelity of at least of 90% must be achieved.

APPENDIX H: Assent Form (8-17 years)

ASSENT FORM (8-17 years)
Project Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.
Principal Investigators: Jorem EmmillianAwadu

Study ID: _____

ASSENT FORM

The investigator will read this assent form to the child at the time of the assessment.
Investigators:

Principal Investigators: Jorem Emmillian Awadu, MA, ABD, Michigan State University
Co-Investigators: Gloria Lee, PhD, Michigan State University, Martin Volker, PhD, Michigan State University & Michael Boivin, PhD, Michigan State University

Purpose:

We are doing a study to learn about children like you; who attend your school (or who come to Butabiika hospital. We would love to know how you think, learn, and make friends. What we need to do is to ask your parents some questions about you. This will be done only once with your parents.
If you do or do not want me to talk to your parents, it is your choice. No one will be upset with you. The information your parents share with me will not be shared with others apart from only those doing the study.

A description of sponsors of the research project and the organizational affiliation of the researchers:

Sponsor

Michigan State University.

Organizational affiliation of the researchers

1. Michigan State University.

Who will participate in the study?

If you agree for your parents to participate, they will be one of up to 100 other parents with children of ages 4 to 18 years.

ASSENT FORM (8-17 years)
Project Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.
Principal Investigators: Jorem EmmillianAwadu

Study ID: _____

How long each will be required to be active in the study:

Your parents will be required to respond to some questions about you for about one hour. This will be done only once.

Risks or discomforts:

1) The risk of this study to you is that you might have to wait for an hour as your parent responds to the questions, which will make you tired. This will occur only if you come along with your parent on the day she or he is to respond to the questions.

Benefits:

The benefits of the study are:

1. If we find that you may need further assessment during our interview with your parents, we shall advise them to take you to better-trained people for further help.

Confidentiality:

1) Rigorous efforts will be made to keep your personal information confidential. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Your identity will be held in confidence in reports in which the study may be published and in databases in which results may be stored.
2) Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, Makerere University School of Health Sciences Research and Ethics Committee or its designees and the study sponsor, Michigan State University who may need to access your research records.

Alternatives:

ASSENT FORM (8-17 years)
Project Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.
Principal Investigators: Jorem EmmillianAwadu

Study ID: _____

Participation is voluntary, you may choose not to participate at all, or you may refuse to participate in certain procedures or answer certain questions or discontinue your participation at any time. This will not affect in any way the services you receive within the hospital or schools now or in the future.

Compensation and reimbursement for participation in the study:

Yourparent will receive 10,000 Ugandanshillingsas a tokenofappreciationforparticipating in thisstudy.

Questions about the study:

For questions about the study or a research-related, please contact the researcher Jorem Awadu at 0771-998931.

Questions about participants rights:

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Makerere University School of Health Sciences Research and Ethics Committee chairperson (Dr. Paul Kutyabami) on telephone +256-772404970, or SHSREC office line (256)-0200903786. You can also contact Uganda National Council of Science and Technology on telephone (0414705500).

Dissemination of results:

Results from this research may be presented in publications and meetings but your name will not be identified.

Ethical approval:

The Makerere University School of Medicine Research and Ethics Committee (SOMREC), and the Uganda National Council of Science and Technology (UNCST) have approved this study.

Statement of Assent:

NOTE: Do not sign this consent form if it does not have an IRB approval stamp, or if the date has lapsed

MAKERERE UNIVERSITY
SCHOOL OF HEALTH SCIENCES
APPROVED
VALID UNTIL
★ 25 JUN 2020 ★
RESEARCH & ETHICS COMMITTEE
P. O. BOX 7072, KAMPALA

MAKERERE UNIVERSITY
SCHOOL OF HEALTH SCIENCES
AMENDMENT APPROVED
★ 24 JUL 2019 ★
INSTITUTIONAL REVIEW BOARD
P.O. BOX 7072, KAMPALA

ASSENT FORM (8-17 years)
Project Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.
Principal Investigators: Jorem EmmillianAwadu

Study ID: _____

..... has described to me what is going to be done, the risks, the benefits involved and my rights regarding this study. I understand that the decision for parents to participate or not to do so in this study will not alter my usual services. In the use of this information, my identity will be concealed. I understand that by signing this form, I do not waive any of my legal rights but merely indicate that I have been informed about the research study in which I am voluntarily agreeing for my parents to participate. A copy of this form will be provided to me.

If you agree, circle "YES", If you do not agree, circle "NO"

☐ **YES** ☐ **NO**

_____
Name of child (Print)

_____ Date_____
Name of child (Signature or mark of assent)

_____
Name of person obtaining assent (print)

_____
Signature of person obtaining assent Date_____

NOTE: Do not sign this consent form if it does not have an IRB approval stamp, or if the date has lapsed

MAKERERE UNIVERSITY SCHOOL OF HEALTH SCIENCES APPROVED VALID UNTIL ★ 25 JUN 2020 ★ RESEARCH & ETHICS COMMITTEE P. O. BOX 7072, KAMPALA

MAKERERE UNIVERSITY SCHOOL OF HEALTH SCIENCES AMENDMENT APPROVED ★ 24 JUL 2019 ★ INSTITUTIONAL REVIEW BOARD P.O. BOX 7072, KAMPALA

ASSENT FORM (8-17 years)

Project Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.

Principal Investigators: Jorem EmmillianAwadu

Study ID: _____

*If the child is unable to read and/or write, a witness should be present during the informed assent discussion. By signing the assent form, the witness attests that the information in the assent form and any other written information were explained to and understood by the child and the child freely gave that informed assent.

Name of Person Witnessing Assent (Printed)_____

Signature of Person Witnessing Assent _____ Date/Time

APPENDIX I: Consent Form

CONSENT FORM

Project Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.

Principal Investigators: Jorem Emmillian Awadu

Study number: _____

CONSENT FORM:

A. Introduction: We ask that you read this form and ask any questions you may have before agreeing to be in the study. We expect that this process will take about15 minutes to complete.

Investigators:
Principle investigator:
Jorem Emmillian Awadu (ABD)
Department of Counseling, Educational Psychology and Special Education
Michigan State University
Email: awadujor@msu.edu
Telephone: +256 771998931

Co-investigators:
Dr. Gloria Lee
Associate Professor
Department of Counseling, Educational Psychology and Special Education
Michigan State University
Email: leekalai@msu.edu
Telephone: +1 517-432-3623

Dr. Martin Volker
Associate Professor
Department of Counseling, Educational Psychology and Special Education
Michigan State University
Email: volkerma@msu.edu
Telephone: +1 517-432-1547

Dr. Michael Boivin
Professor
Department of Psychiatry
Michigan State University
Email: boivin@msu.edu
Telephone: +1 (517) 884-0281

Information about the study and why it is being carried out:

We know that one out of 145 children have autism worldwide and that most parents find out that their child has autism between two and three years of age. Other children especially those living in rural areas with a few well-trained professionals in autism could be diagnosed at a much later time such as seven or more years. Helping the child gain some of the skills like speaking, toileting, bathing,

socialbehavior and others therefore start late and yet research shows that the child improves better when rehabilitation is done early. This study is therefore to help find a way to identify children who may have autism so that they can be sent to professionals with more knowledge on autism for a diagnosis. This will enable the children to start the rehabilitation early and increase on the chance that they can gain some of the skills for a better adulthood.

Study sponsor:
This study is sponsored through a research fellowship from the College of Education,department of Counseling, Educational Psychology and Special Education in Michigan State University. All investigators except Jorem Emmillian Awadu are faculty of Michigan State University. Jorem Emmillian Awadu is a doctoral student in the department that funded this research.

B. Study Purpose:
The purpose of this study is to examine how well the Social Responsiveness Scale, Social Communication Questionnaire, and the 23-Q help us identify children at-risk ofautism in Uganda.

C. Study Procedures:
If you choose to participate in this study, you will be requested to respond to a questionnaire related to you and your child. You as the parent or caregiver, will be asked to complete three surveys, read by a research assistant, about the behaviours of your child. In addition, the researcher may also request a copy of your child's medical records.

D. Study Participants and sites
The study participants will be parents or caretakers of children with autism (aged 4 to 18 years) and living within Kampala,Wakiso and Entebbe districts.

D. Risks of Study Participation

We anticipate that the completion of these surveys will pose **less than minimal risk,** as these surveys are common instruments for professionals to use to do diagnosis for autism. However, the three likely risks that may occur include the following.

1) There is a likely risk that the surveys can be long and may cause you to be tired. You can take breaks when you feel tired.

2) The items from the surveys may make you feel upset because these items ask about autistic symptoms that your child may have. The research assistant and the professionals that your child is seeking help from are equipped to assist you to process your feelings.

3) There is minimalrisk of the release of information from your child's diagnosis and or health-related records. First, no identifiable information about you and your child will be on any of the forms. You and your child will be assigned a study number. The research assistant and professionals are also trained to handle all patient and research records confidentially. All research materials will be kept safe and locked, and will not be shared with anyone else who is not part of the study team. All research materials will be destroyed three years after the completion of this study.

CONSENT FORM

Project Title: Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.

Principal Investigators: Jorem Emmillian Awadu

E. Benefits of Study Participation

1. There are no immediate benefits for participating in the current study. However, your participation will help to improve the identification of children at risk of Autism.
2. We also hope that children identified as at risk of autism will be in position to access better trained professionals who can offer a final diagnosis and recommend rehabilitation services.

F. Alternatives to Study Participation

Participation is voluntary, you may choose not to participate at all, or you may refuse to participate in certain procedures or answer certain questions or discontinue your participation at any time without consequence. This will not affect in any way the continued support services you receive within the hospital or schools now or in the future.

G. Study Costs/Compensation

You will not incur any costs in participating in this study. You will receive 10,000 Ugandan shillings as a token of appreciation for participating in this study.

H. Research Related Injury

There is no foreseeable research related injury.

I. Confidentiality

The results of this study will be kept strictly confidential, and used only for research purposes. My identity will be concealed in as far as the law allows. My name will not appear anywhere on the coded forms with the information. Paper and computer records will be kept under lock and key and with password protection respectively.

The interviewer has discussed this information with me and offered to answer my questions. For any further questions, I may contact the Chairperson of the School of Health Sciences Research and Ethics Committee (MakSHSREC) on (+256) 772-404970 / (+256) 0200903786 / or Uganda National Council of Sciences and Technology. Tel: (+256)-041-4705500).

J. Contacts and Questions

If you have concerns or questions about this study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researcher**Jorem Emmillian Awadu at awadujor@msu.edu**(mobile: +256 771998931).

K. Participant Rights

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Makerere University School of Health Sciences Research and Ethics Committee chairperson (Dr. Paul Kutyabami) on telephone +256-772404970, or SHSREC office line (256)-0200903786. You can also contact Uganda National Council of Science and Technology on telephone (0414705500).

L. Feedback on Study findings and progress of the study

Parents whose children are found to be at-risk of autism will be referred further diagnostic assessment by advanced professionals. The final results from the study will be shared with all the study

NOTE: Do not sign this consent form if it does not have an IRB approval stamp, or if the date has lapsed

MAKERERE UNIVERSITY SCHOOL OF HEALTH SCIENCES APPROVED VALID UNTIL 25 JUN 2020 RESEARCH & ETHICS COMMITTEE P. O. BOX 7072, KAMPALA

stakeholders including; Makerere School of Health Sciences, Butabika and Mulago hospitals as well as principals of the participating schools.

L. Statement of Consent

..... has described to me what is going to be done, the risks, the benefits involved and my rights regarding this study. I understand that my decision to participate in this study will not alter my usual medical care. In the use of this information, my identity will be concealed. I am aware that I may withdraw at anytime. I understand that by signing this form, I do not waive any of my legal rights but merely indicate that I have been informed about the research study in which I am voluntarily agreeing to participate. A copy of this form will be provided to me.

Name .....

Signature of parent/guardian for minors .....

Date .....

Name.....

Signature of witness (if applicable)..... .....

Date.....

Name .....

Signature of interviewer/Person obtaining informed consent

.....

Date .....

NOTE: Do not sign this consent form if it does not have an IRB approval stamp, or if the date has lapsed

MAKERERE UNIVERSITY
SCHOOL OF HEALTH SCIENCES
APPROVED
VALID UNTIL
25 JUN 2020
RESEARCH & ETHICS COMMITTEE
P. O. BOX 7072, KAMPALA

APPENDIX J: Demographic Form

Validation of autism assessments: Comparison of the Social Communication Questionnaire, Social Responsiveness Scale and 23-Q with DSM- 5 in assessing for autism spectrum disorder (ASD) in Uganda.

DEMOGRAPHICS 1
ID _____

Directions: Please complete the following form as completely as possible. Please answer all questions truthfully. This information is being collected for research purposes only. Your information will be kept confidential and will not be shared in any way by the research team.

Child Information

1. Please indicate your child's gender: M F
2. What is your child's birthdate? _____
3. What is your child's Ethnicity?

   ___ African
   ___ European
   ___ Asian
   ___ Other, please specify: _____
4. At what age was your child diagnosed with Autism Spectrum Disorder?_____
5. My child was first diagnosed with autism by a:

   ___ Clinical psychologist
   ___ School psychologist
   ___ Medical doctor
   ___ Other, please specify: _____
6. Does your child have any other identified delays or disabilities (besides Autism Spectrum Disorder), such as ADHD, speech delay, or other special needs? Yes No
7. If you answered "yes" to #6, please indicate what other delays/disability(ies) your child has been identified as having.

   _____
8. What, if any, specialized services is your child currently receiving? (e.g., speech therapy, private counseling, special education)?

   _____

   _____
9. How muchin Ugandan shillings do pay for the specialized services your child is currently receiving? (e.g., speech therapy, private counseling, special education)?

1

DEMOGRAPHICS 2
ID _____

_____

_____

10. What, if any, specialized services have been provided to your child in the past?

_____

_____

_____

_____

Parent Information

11. What is your birthdate? _____
12. Which district do you currently live in? _____
13. What is your tribe? _____
14. What describes your highest level of education?

_____ None
_____ Primary school (P.1-P.7)
_____ Secondary school (Ordinary/Advanced Levels).**Please circle one.**
_____ Trade/Technical/Vocational Training
_____ Diploma
_____ Bachelors Degree
_____ Graduate Degree (master's level or higher)
_____ Don't Know
_____ Other, please specify: _____

15. Please indicate your marital status:

_____Single
_____Married
_____Living with domestic partner
_____Divorced
_____Widowed
_____Other, please specify:_____

16. Is the child's mother employed? Yes No

DEMOGRAPHICS 3
ID _____

17. What is the occupation of the mother_____
18. Is the child's father employed? Yes No
19. What is the occupation of the father?_____

20. What is the combined estimated monthly family earnings from allsources_____

23 Questions Questionnaire

Initial Screening Form

Child ID number: _____ Assessment Date, Month, and Year: _____

Age _____yrs_____months Assessment Performed By: _____

Sex *(circle one)* M F Respondent's relationship to Child: _____
Mother=1; Father=2; Grandmother=3; Sibling=4; other relative=5; other=6

Please circle the respondent's answer following each question.

Cerebral Palsy

1. Compared with other children, did the child have any serious delay in sitting, standing, or walking? YES* No DK

2. Does the child have difficulty in walking or moving his/her arms or does he/she have weakness and/or stiffness in the arms or legs? YES* No DK

Visual Impairment

3. Compared with other children, does the child have difficulty seeing either in the day time or at night? YES* No DK
4. In school going children ask: Does the child read well what is written on the blackboard when sitting at the back of the class? YES No* DK

Hearing Impairment

5. Does the child appear to have difficulty hearing? YES* No DK
6. Does the child often ask you to repeat what you have said? YES* No DK

Mental Retardation

7. When you tell the child to do something, does he/she seem to understand what you are saying? YES No* DK
8. Does the child learn to do things like other children his/her age? YES No* DK
9. Compared with other children of his/her age, does the child appear in any way mentally backward, dull, or slow? YES* No DK

Epilepsy

10. Does the child sometimes have fits, become rigid, or lose consciousness? YES* No DK
11. Does the child have episodes of staring when you can not get their attention by talking to them or touching them lightly? YES* No DK

Speech and Language Impairment

12. Does the child speak at all (can he/she make himself/herself understood in words; can he/she say any recognizable words? YES No* DK
13. **For 3-9 year-olds ask**: Is the child's speech in any way different from normal (not clear enough to be understood by people other than his/her immediate family? YES* No DK

    For 2 year-olds ask: Can he/she name at least one object (for example: an animal, a toy, a cup, a spoon)? YES No* DK

1

Autism Spectrum Disorders

14. Does the child shave difficulty making and maintaining eye contact?	YES*	No	DK
15. Does the child cry or get upset if you do not do particular routines the same way every day like using the same plate/cup to serve his/her food/drink, letting him/her sit on a particular stool/chair in the house?	YES*	No	DK
16. Does the child take an interest in playing with other children?	YES	No*	DK
17. Does the child usually turn to look at you when you call his/her name?	YES	No*	DK
18. Does the child repeat phrases over and over exactly as they were said or heard from someone (or on the radio)?	YES*	No	DK

<u>Additional questions for children <5 years of age</u>

19. Does the child engage in pretend play like 'Mama ne Tata', driving, cooking?	YES	No*	DK
20. Does the child usually use his/her index finger to point, to indicate interest in something?	YES	No*	DK
21. Does the child often bring objects over to you (parent) to show you something?	YES	No*	DK
22. Does the child imitate you? (e.g., if you make a face, will the child imitate it?)	YES	No*	DK
23. If you point at a toy or a person across the room, does the child look at it?	YES	No*	DK

Screening result is positive if any one or more of the responses with an asterisk (*) is checked

REFERENCES

REFERENCES

Abubakar, A., Ssewanyana, D., & Newton, C. R. (2016). A Systematic Review of Research on Autism Spectrum Disorders in Sub-Saharan Africa. *Behavioural Neurology*, *2016*, 1–14. https://doi.org/10.1155/2016/3501910

Akoury-Dirani, L., Alameddine, M., & Salamoun, M. (2013). Validation of the Lebanese Childhood Autism Rating Scale – Second Edition – High Functioning Version. *Research in Autism Spectrum Disorders*, *7*(11), 1332–1338. https://doi.org/10.1016/j.rasd.2013.08.001

Aldosari, M., Fombonne, E., Aldhalaan, H., Ouda, M., Elhag, S., Alshammari, H., Ghazal, I., Alsaleh, A., Alqadoumi, T., Thomson, R., Al Khasawneh, M., Tolefat, M., & Alshaban, F. (2019). Validation of the Arabic version of the Social Communication Questionnaire. *Autism*, *23*(7), 1655–1662. https://doi.org/10.1177/1362361318816065

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Avcil, S., Baykara, B., Baydur, H., Munir, K. M., & Inal Emiroglu, N. (2014). The Validity and Reliability of the Turkish Version of the Social Communication Questionnaire. *Turkish Journal of Psychiatry*. https://doi.org/10.5080/u7298

Bakare, M., & Munir, K. (2011). Autism spectrum disorders (ASD) in Africa: a perspective. *African Journal of Psychiatry*, *14*(3). https://doi.org/10.4314/ajpsy.v14i3.3

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (n.d.). *The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians*. 14.

Bello-Mojeed, M. A., Omigbodun, O. O., Bakare, M. O., & Adewuya, A. O. (2017). Pattern of impairments and late diagnosis of autism spectrum disorder among a sub-Saharan African clinical population of children in Nigeria. *Global Mental Health*, *4*. https://doi.org/10.1017/gmh.2016.30

Bernal, G. E., & Domenech Rodríguez, M. M. (2012). *Cultural adaptations: Tools for evidence-based practice with diverse populations* (pp. xix-307). American Psychological Association.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: diagnostic validity. *The British Journal of Psychiatry*, *175*(5), 444–451.

Bölte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Research*, *1*(6), 354–363. https://doi.org/10.1002/aur.49

Bölte, S., Westerwald, E., Holtmann, M., Freitag, C., & Poustka, F. (2011). Autistic Traits and Autism Spectrum Disorders: The Clinical Validity of Two Measures Presuming a Continuum of Social Communication Skills. *Journal of Autism and Developmental Disorders*, *41*(1), 66–72. https://doi.org/10.1007/s10803-010-1024-9

Bryson, S. E., Rogers, S. J., & Fombonne, E. (2003). Autism Spectrum Disorders: Early Detection, Intervention, Education, and Psychopharmacological Management. *The Canadian Journal of Psychiatry*, *48*(8), 506–516. https://doi.org/10.1177/070674370304800802

Cen, C.-Q., Liang, Y.-Y., Chen, Q.-R., Chen, K.-Y., Deng, H.-Z., Chen, B.-Y., & Zou, X.-B. (2017). Investigating the validation of the Chinese Mandarin version of the Social Responsiveness Scale in a Mainland China child population. *BMC Psychiatry*, *17*(1), 51. https://doi.org/10.1186/s12888-016-1185-y

Chlebowski, C., Green, J. A., Barton, M. L., & Fein, D. (2010). Using the Childhood Autism Rating Scale to Diagnose Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *40*(7), 787–799. https://doi.org/10.1007/s10803-009-0926-x

Chojnicka, I., & Pisula, E. (2019). Cross-Cultural Validation of the Polish Version of the ADI-R, Including New Algorithms for Toddlers and Young Preschoolers. *Child Psychiatry & Human Development*. https://doi.org/10.1007/s10578-018-00865-2

Cholemkery, H., Kitzerow, J., Rohrmann, S., & Freitag, C. M. (2014). Validity of the social responsiveness scale to differentiate between autism spectrum disorders and disruptive behaviour disorders. *European Child & Adolescent Psychiatry*, *23*(2), 81–93. https://doi.org/10.1007/s00787-013-0427-5

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook, E. H., Leventhal, B. L., & Lord, C. (2007). Between a ROC and a hard place: decision making and making decisions about using the SCQ. *Journal of Child Psychology and Psychiatry*, *48*(9), 932–940. https://doi.org/10.1111/j.1469-7610.2007.01762.x

Contantino, J. N., & Gruber, C. P. (2005). Social Responsiveness Scale (SRS) [Manual]. Los Angeles, CA: Western Psychological Services.

Constantino, J. N., & Gruber. C. P. (2012). Social Responsiveness Scale, Second Edition (SRS-2) [Manual]. Torrance, CA: Western Psychological Services.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., … Varley, J. (2010). Randomized, Controlled Trial of an Intervention for Toddlers With Autism: The Early Start Denver Model. *Pediatrics*, *125*(1), e17–e23. https://doi.org/10.1542/peds.2009-0958

Durkin, M. (1995). Evaluating a ten questions screen for childhood disability: Reliability and internal structure in different cultures. *Journal of Clinical Epidemiology*, *48*(5), 657–666. https://doi.org/10.1016/0895-4356(94)00163-K

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., … Shih, A. (2015). Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide: Enhancing Autism Research and Services Worldwide. *Autism Research*, *8*(5), 473–476. https://doi.org/10.1002/aur.1575

Fombonne, E., Marcin, C., Bruno, R., Tinoco, C. M., & Marquez, C. D. (2012). Screening for Autism in Mexico: Screening for autism in Mexico. *Autism Research*, *5*(3), 180–189. https://doi.org/10.1002/aur.1235

Franz, L., Chambers, N., von Isenburg, M., & de Vries, P. J. (2017). Autism spectrum disorder in sub-saharan africa: A comprehensive scoping review: Autism spectrum disorder in sub-Saharan Africa. *Autism Research*, *10*(5), 723–749. https://doi.org/10.1002/aur.1766

Gau, S. S.-F., Lee, C.-M., Lai, M.-C., Chiu, Y.-N., Huang, Y.-F., Kao, J.-D., & Wu, Y.-Y. (2011). Psychometric properties of the Chinese version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders*, *5*(2), 809–818. https://doi.org/10.1016/j.rasd.2010.09.010

Gilliam, J. E. (2014). Gilliam Autism Rating Scale (3rd ed.). Austin, TX: Pro-Ed.

Gona, J. K., Newton, C. R., Rimba, K., Mapenzi, R., Kihara, M., Van de Vijver, F. J. R., & Abubakar, A. (2015). Parents' and Professionals' Perceptions on Causes and Treatment Options for Autism Spectrum Disorders (ASD) in a Multicultural Context on the Kenyan Coast. *PLOS ONE*, *10*(8), e0132729. https://doi.org/10.1371/journal.pone.0132729

Harrison, A. J., Zimak, E. H., Sheinkopf, S. J., Manji, K. P., & Morrow, E. M. (2014). Observation-centered Approach to ASD Assessment in Tanzania. *Intellectual and Developmental Disabilities*, *52*(5), 330–347. https://doi.org/10.1352/1934-9556-52.5.330

Jackson, L. S., Little, S. G., & Akin-Little, A. (2013). The Spanish adaptation of the Gilliam Autism Rating Scale-2: Translation and psychometric analysis. *Research in Autism Spectrum Disorders*, *7*(9), 1160–1167. https://doi.org/10.1016/j.rasd.2013.06.005

Kakkar, J., & Srivastava, P. (n.d.). *Challenges and Coping among Parents having Children with Autism Spectrum Disorder*. 10

Kakooza-Mwesige, A., Ssebyala, K., Karamagi, C., Kiguli, S., Smith, K., Anderson, M. C., … Grether, J. K. (2014). Adaptation of the "ten questions" to screen for autism and other neurodevelopmental disorders in Uganda. *Autism*, *18*(4), 447–457. https://doi.org/10.1177/1362361313475848

Kamp-Becker, I., Albertowski, K., Becker, J., Ghahreman, M., Langmann, A., Mingebach, T., … Stroth, S. (2018). Diagnostic accuracy of the ADOS and ADOS-2 in clinical practice. *European Child & Adolescent Psychiatry*, *27*(9), 1193–1207. https://doi.org/10.1007/s00787-018-1143-y

Karren, B. C. (2017). A Test Review: Gilliam, J. E. (2014). *Gilliam Autism Rating Scale–Third Edition (GARS-3)*. *Journal of Psychoeducational Assessment*, *35*(3), 342–346. https://doi.org/10.1177/0734282916635465

Klin, A., & Volkmar, F. R. (2005). Editorial Preface. *Journal of Autism and Developmental Disorders*, *35*(2), 141–143. https://doi.org/10.1007/s10803-004-1991-9

Lai, M.-C. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *ADOLESCENT PSYCHIATRY*, *54*(1), 14.

Lecavalier, L. (2005). An Evaluation of the Gilliam Autism Rating Scale. *Journal of Autism and Developmental Disorders*, *35*(6), 795–805. https://doi.org/10.1007/s10803-005-0025-6

Lee, G. K., Lopata, C., Volker, M. A., Thomeer, M. L., Nida, R. E., Toomey, J. A., … Smerbeck, A. M. (2009). Health-Related Quality of Life of Parents of Children With High-Functioning Autism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities*, *24*(4), 227–239. https://doi.org/10.1177/1088357609347371

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., … Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.

Lord C, Rutter M, Dilavore PC, Risi S, Gotham K, Bishop SL (2012). Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (part 1) modules 1-4. Western Psychological Services, Torrance, CA

Maenner, M. J., Shaw, K. A., Baio, J., EdS1, Washington, A., Patrick, M., DiRienzo, M., Christensen, D. L., Wiggins, L. D., Pettygrove, S., Andrews, J. G., Lopez, M., Hudson, A., Baroud, T., Schwenk, Y., White, T., Rosenberg, C. R., Lee, L. C., Harrington, R. A., Huston, M., … Dietz, P. M. (2020). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016. *Morbidity and mortality weekly report. Surveillance summaries (Washington, D.C. : 2002)*, *69*(4), 1–12. https://doi.org/10.15585/mmwr.ss6904a1

Magaña, S., & Smith, L. E. (2013). The Use of the Autism Diagnostic Interview-Revised with a Latino Population of Adolescents and Adults with Autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1098–1105. https://doi.org/10.1007/s10803-012-1652-3

Magaña, S., & Vanegas, S. B. (2017). Diagnostic Utility of the ADI-R and DSM-5 in the Assessment of Latino Children and Adolescents. *Journal of Autism and Developmental Disorders*, *47*(5), 1278–1287. https://doi.org/10.1007/s10803-017-3043-2

Marlow, M., Servili, C., & Tomlinson, M. (2019). A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low- and middle-income countries: Marlow et al./A review of screening tools for autism and developmental delay. *Autism Research*, *12*(2), 176–199. https://doi.org/10.1002/aur.2033

Mazurek, M. O., Curran, A., Burnette, C., & Sohl, K. (2019). ECHO Autism STAT: Accelerating Early Access to Autism Diagnosis. *Journal of Autism and Developmental Disorders*, *49*(1), 127–137. https://doi.org/10.1007/s10803-018-3696-5

Miller, J. N., & Ozonoff, S. (1997). Did Asperger's Cases Have Asperger Disorder? A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(2), 247–251. https://doi.org/10.1111/j.1469-7610.1997.tb02354.x

Mullan, F., Frehywot, S., Omaswa, F., Buch, E., Chen, C., Greysen, S. R., … Neusy, A.-J. (2011). Medical schools in sub-Saharan Africa. *The Lancet*, *377*(9771), 1113–1121. https://doi.org/10.1016/S0140-6736(10)61961-7

Nguyen, P. H., Ocansey, M. E., Miller, M., Le, D. T. K., Schmidt, R. J., & Prado, E. L. (2019). The reliability and validity of the social responsiveness scale to measure autism symptomology in Vietnamese children. Autism Research, 12(11), 1706–1718. https://doi.org/10.1002/aur.2179

Nordin, V., Gillberg, C., & Nyden, A. (n.d.). *The Swedish Version of the Childhood Autism Rating Scale in a Clinical Setting*. 7.

Norris, M., & Lecavalier, L. (2010). Screening Accuracy of Level 2 Autism Spectrum Disorder Rating Scales: A Review of Selected Instruments. *Autism*, *14*(4), 263–284. https://doi.org/10.1177/1362361309348071

Presmanes Hill, A., Zuckerman, K., & Fombonne, E. (2015). Epidemiology of Autism Spectrum Disorders. In M. de los A. Robinson-Agramonte (Ed.), *Translational Approaches to Autism Spectrum Disorder* (pp. 13–38). Retrieved from http://link.springer.com/10.1007/978-3-319-16321-5_2

Rutter, M., Bailey, A., & Lord, C, (2003). The Social Communication Questionnaire (SCQ) [Manual]. Torrance, CA: Western Psychological Services.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, *10*(1), 91–103. https://doi.org/10.1007/BF02408436

Smith, L., Malcolm-Smith, S., & de Vries, P. J. (2017). Translation and cultural appropriateness of the Autism Diagnostic Observation Schedule-2 in Afrikaans. *Autism*, *21*(5), 552–563. https://doi.org/10.1177/1362361316648469

South, M., Williams, B. J., McMahon, W. M., Owley, T., Filipek, P. A., Shernoff, E., … Ozonoff, S. (2002). *Utility of the Gilliam Autism Rating Scale in Research and Clinical Populations*. 7.

Sturm, A., Kuhfeld, M., Kasari, C., & McCracken, J. T. (2017). Development and validation of an item response theory-based Social Responsiveness Scale short form. *Journal of Child Psychology and Psychiatry*, *58*(9), 1053–1061. https://doi.org/10.1111/jcpp.12731

Tehrani-Doost, M., Shahrivar, Z., Torabi, N., Ansari, S., Haji-Esmaeelzadeh, M., & Saeed-Ahmadi, S. (2018). Cross-Cultural Validation and Normative Data of the Social Responsiveness Scale in a Group of Iranian General Child Population. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-018-3773-9

Tsuchiya, K. J., Matsumoto, K., Yagi, A., Inada, N., Kuroda, M., Inokuchi, E., … Takei, N. (2013). Reliability and Validity of Autism Diagnostic Interview-Revised, Japanese Version. *Journal of Autism and Developmental Disorders*, *43*(3), 643–662. https://doi.org/10.1007/s10803-012-1606-9

Vanegas, S. B., Magaña, S., Morales, M., & McNamara, E. (2016). Clinical Validity of the ADI-R in a US-Based Latino Population. *Journal of Autism and Developmental Disorders*, *46*(5), 1623–1635. https://doi.org/10.1007/s10803-015-2690-4

Vaughan, C. A. (2011). Test Review: E. Schopler, M. E. Van Bourgondien, G. J. Wellman, & S. R. Love Childhood Autism Rating Scale (2nd ed.). Los Angeles, CA: Western Psychological Services, 2010. *Journal of Psychoeducational Assessment*, *29*(5), 489–493. https://doi.org/10.1177/0734282911400873

Volker, M. A., & Lopata, C. (2008). Autism: A review of biological bases, assessment, and intervention. *School Psychology Quarterly*, *23*(2), 258–270. https://doi.org/10.1037/1045-3830.23.2.258

Wang, J., Lee, L.-C., Chen, Y.-S., & Hsu, J.-W. (2012). Assessing Autistic Traits in a Taiwan Preschool Population: Cross-Cultural Validation of the Social Responsiveness Scale (SRS). *Journal of Autism and Developmental Disorders*, *42*(11), 2450–2459. https://doi.org/10.1007/s10803-012-1499-7

Wing, L. (1986). Clarification on Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *16*(4), 513–515. https://doi.org/10.1007/BF01531716

Witwer, A. N., & Lecavalier, L. (2007). Autism screening tools: An evaluation of the Social Communication Questionnaire and the Developmental Behaviour Checklist–Autism Screening Algorithm. *Journal of Intellectual & Developmental Disability*, *32*(3), 179–187. https://doi.org/10.1080/13668250701604776

Woodbury-Smith, M., Klin, A., & Volkmar, F. (2005). Asperger?s Syndrome: A Comparison of Clinical Diagnoses and Those Made According to the ICD-10 and DSM-IV. *Journal of Autism and Developmental Disorders*, *35*(2), 235–240. https://doi.org/10.1007/s10803-004-2002-x

Woolfenden, S., Sarkozy, V., Ridley, G., & Williams, K. (2012). A systematic review of the diagnostic stability of Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, *6*(1), 345–354. https://doi.org/10.1016/j.rasd.2011.06.008

Zander, E., Willfors, C., Berggren, S., Choque-Olsson, N., Coco, C., Elmund, A., … Bölte, S. (2016). The objectivity of the Autism Diagnostic Observation Schedule (ADOS) in naturalistic clinical settings. *European Child & Adolescent Psychiatry*, *25*(7), 769–780. https://doi.org/10.1007/s00787-015-0793-2

Zarokanellou, V., Kolaitis, G., Vlassopoulos, M., & Papanikolaou, K. (2017). Brief report: A pilot study of the validity and reliability of the Greek version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders*, *38*, 1–5. https://doi.org/10.1016/j.rasd.2017.03.001