Screening for Autism and Intellectual Disabilities amongst the Nigerian Adolescent Population

Eziafakaku U Nwokolo

Tizard Centre

School of Social Policy, Sociology and Social Research

University of Kent

A thesis submitted for the degree of

Doctor of Philosophy

Intellectual and Developmental Disabilities

February 2023

Words 87,822

Acknowledgements

Where do I start? So many people to thank and acknowledge but let me specifically thank the following people.

I want to thank my supervisors, Prof Peter Langdon and Prof Glynis Murphy, for their invaluable guidance. For their ideas, suggestions, edits (red pen), feedback and laughter. When I kicked against some of the frameworks (I honestly did not know how herculean a PhD was!!), you remained patiently dogged and doggedly patient. For that, I am grateful. The pandemic caught us at a bad time, and I was tugged in so many different directions that I cannot even begin to write about it, but you remained rock solid and helped me navigate all the rough terrains. Thanks for all your patience, listening ears and standing up to the status quo.

Thanks to my children, Nmesomachi, Nulia and Nodebenna, for spurring me on and believing I could do it. For understanding when my visits were short and when I could not be there even when I wanted to or when I was there but buried between articles. God bless you, girls. To my husband, who planted the idea of doing a PhD midway into my master's programme, for his unwavering support over the past five years of my never-ending studying. All those nights of sleeping on the couch while waiting as I scoured over articles and pounded away on my keyboard, bless you. Thank you, mummy, for all your prayers, love, support and for checking in on me. Now we can gist.

God bless all the participants, without whom there would be no research. May you find true help when you need it.

Thank you to all my extended family and friends who understood when I was not there for big occasions. I am finally done with the PhD.

Finally, I am grateful to God Almighty for life, strength, and, most importantly, grace during trying times.

Dedicated to the loving memory of my late father, Dr Golie Tabugbo Onwuka, aka GT, who set the tune for my siblings and me to walk and do life without any fears. Thanks, Nna, as we fondly called you; you would have been proud of me.

Abstract

Autism and intellectual disabilities awareness is increasing globally, perhaps due to the availability of technology and global migration. The ability to detect these disabilities in the West continues to grow as new screening tools are developed and previous screening and diagnostic methods are revised. This is not the case in Africa and Nigeria specifically. The inability to use standardised measures in assessing for these disabilities poses a challenge and impedes the ability to provide intervention for the affected population. While there are readily available tools online for screening autism in early years, there is no evidence of their cross-cultural validation in most African countries and Nigeria especially. There is also limited evidence for the availability and validation of screening tools for adolescents globally. Therefore, persons who are suspected to have autism in later years are not properly assessed. Regarding intellectual disability, there is no screening tool readily available for adolescents. Diagnosing intellectual disability requires a multi-tiered approach but starts with screening. Assessing adolescents with intellectual disability in Africa is also a challenge as there is limited access to measures by professionals. There are very few screening tools available for intellectual disability and even more limited for adolescents.

This thesis sought to identify screening tools for both intellectual disability and autism, that may be culturally appropriate for use within Africa. Where there are non-readily available, the thesis explored the possibility of adapting existing tools for autism and intellectual disability. The focus on both disabilities is due to their often comorbidity.

Study 1 was a systematic review which identified, described, and critically appraised short screening tools for the detection of intellectual disabilities and autism for older children and young adults. The psychometric properties of these tools were then examined for their cultural appropriateness for use within Africa. Six screening tools for intellectual disabilities and twelve for autism were identified and appraised using the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines. Two screening tools each for autism and intellectual disability appeared appropriate for validation for use within African nations.

Study 2 was a focus group meeting which employed the Nominal Group Technique (NGT). The group examined the identified screening tools for the items' face and content validity and cultural relevance. The group consisted of professionals, parents, and laypersons who

selected two of the four existing screening tools for autism spectrum disorder and intellectual disability for adaptation and use with older children and adolescents in Nigeria.

Study 3a involved the validation of the selected tool for intellectual disability, the English version of the adolescent Screener for Intelligence and Learning Disabilities (SCIL). The psychometric properties of the tool when used with Nigerian adolescents were assessed. There were two hundred and nine adolescents and young people (aged 11 – 26 years) who completed the SCIL and took part in an assessment of their level of general intellectual functioning and adaptive behaviour. The study determined that the SCIL has good psychometric properties when used with Nigerian adolescents.

Study 3b was the validation of the selected tool for autism, the Social Communication Questionnaire (SCQ). The psychometric properties of the tool when used with Nigerian adolescents were assessed. Parents and caregivers of two hundred and five adolescents completed the SCQ Lifetime form while the adolescents were assessed for autism using the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2). Convergent and discriminant validity were examined, along with the sensitivity and specificity of the SCQ in identifying participants with an autism spectrum disorder. The SCQ has good psychometric properties when used with Nigerian adolescents.

Publications

Nwokolo, E. U., Langdon, P. E., & Murphy, G. H. (2022). Screening for Intellectual Disabilities and/or Autism Amongst Older Children and Young Adults: A Systematic Review of Tools for Use in Africa. Review Journal of Autism and Developmental Disorders. Advance online publication. https://doi.org/10.1007/s40489-022-00342-6 (Appendix 19)

Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E., (in press). Using the consensus group method to select the best screening tools for autism and intellectual disability for use with Nigerian adolescents. *Journal of Policy & Practice in Intellectual Disabilities.* (Appendix 20)

Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E., (submitted). Preliminary testing of the English Version of the Screener for Intelligence and Learning Disabilities (SCIL) amongst adolescents in Nigeria. *Journal of Intellectual Disability Research.* (Appendix 21)

Nwokolo, E. U., Murphy, G. H., & Langdon, P. E., (submitted). Validation of the Social Communication Questionnaire (SCQ) amongst Nigerian adolescents. *Autism Research.* (Appendix 22)

Presentations

Nwokolo, E. U., Murphy, G. H., Langdon, P. E., (2021, July 6–8). *Selection of the best screening tools for use with the Nigerian adolescent population: Focus group* [Poster Presentation]. 6th Europe Congress of the International Association for the Scientific Study of Intellectual and Developmental Disabilities. Online Event.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1. Autism and Intellectual Disabilities – Definitions, Terminologies, Description, Theories and Models of Disabilities

This initial chapter will focus on definitions, theories, prevalence rates, descriptions and characteristics of autism spectrum disorder and intellectual disability. The usefulness of the theories and models for screening adolescents with autism and/or intellectual disability will be the focus. Significant theories of autism and models of disability will be explored, and the ways disorders are perceived within the Nigerian context. This chapter will provide the basis for understanding the reasoning behind the state of screening for autism and intellectual disability in Nigeria.

1.1 Autism

1.1.1 Autism's Historical Background

In 1911, the German psychiatrist Eugen Bleuler used the word autism to describe his concept of severe schizophrenia. Bleuler identified and grouped distorted and disorganised mental functions into four classes: ambivalence, affectivity, association, and autism (McGlashan, 2011; Moskowitz & Heim, 2011). Bleuler's autistics were the most severe of the people with schizophrenia; they had no contact with the outside world, lived in a world of their own, were preoccupied with their inner life, and were detached from reality (McGlashan, 2011; Evans, 2013). This terminology was applied to children diagnosed with psychiatric and mental disabilities throughout the late forties and fifties. Leo Kanner, in 1943 identified eleven children who were physically and cognitively healthy but with peculiar psychological disorders. Kanner's children were characterised by obsessiveness, extreme aloneness, limited spontaneity, stereotypy, insistence on sameness and echolalia. In addition, they had excellent rote memory, verbal rituals, were disengaged from their environments and lacked social interaction. Kanner applied the term autism to their condition, referring to these signs as 'inborn autistic disturbances of affective contact' (Kanner, 1943, p. 250). By the sixties, the word autism had evolved further and became applicable to a condition with distinct features from schizophrenia. This evolution of terminology led medical practitioners to view it as a discrete condition requiring further study. Hans Asperger, in his study (Asperger, 1944 & 1991), described similar traits in his participants as Kanner's; the difference lay in Asperger's participant's language development and age of onset. Asperger's children showed autistic traits from age three while Kanner's

were from the first month; Asperger's children communicated, but it was a 'one-way traffic' while Kanner's children had non-functional language (Van Krevelen, 1971).

Currently, autism is understood to be a lifelong disability, distinct from schizophrenia and other neurodevelopmental disorders (Evans, 2013; World Health Organization (WHO), 2020; American Psychiatric Association (APA), 2013), and there have been a number of developments in describing the characteristics of those with autism. Due to the different cultural, professional, and individual perspectives on current terminology, person-first (person/individual with autism) and identity-first (autistic) language will be used interchangeably in this thesis (Crocker & Smith, 2019; Shakes & Cashin, 2019; Anderson-Chavarria, 2021; Buijsman, Begeer & Scheeren, 2022).

1.2 Characteristics of Autism

Several features described by Kanner and Asperger are still essential and remain relevant in understanding autism today. However, the departure from hallucinations, childhood psychosis, and schizophrenia towards a definition as a developmental disorder accompanied by language and communication problems led to the redefinition of autism and inclusion in the International Classification of Diseases, 8th Edition (WHO, 1967) and the Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (APA, 1980). A very influential contribution came from Wing & Gould (1979), who described the core features of autism, labelling them as the triad of impairments. Deficits in the social imagination, social interaction and social communication formed the triad. Recently, the International Classification of Diseases, 11th Edition ([ICD-11], WHO, 2020) and the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition ([DSM-5] APA, 2013) redefined the specifiers for autism spectrum disorder (Ousley & Cermak, 2014; Rapoport, Chavez, Greenstein, Addington, & Gogtay, 2009). The DSM-5 and ICD-11 have also reduced the dimensions of autism to a dyad– social communication and social interaction impairments and repetitive and restricted behaviours.

To meet the autism classification under the ICD-11, individuals must meet the following core features:

1. persistent deficits in reciprocal social interaction and social communication
2. persistent restricted, repetitive, and inflexible patterns of behaviour, interests, or activities

3. the onset of the disorder should be typically in early childhood, although there may be some delays in the presentation of symptoms.

Under the DSM-5, to be classified as autistic, the following features must be met:

1. persistent deficits in social communication and social interaction across multiple contexts and the
2. presence of restricted, repetitive patterns of behaviour, interests, or activities.

Although autistic individuals present with these core features, autism is heterogeneous. The spectrum is recognised as a continuum, and each autistic person presents differently (Lord & Jones, 2012; Grzadzinski, Huerta, & Lord, 2013). Because of the heterogeneity, functioning, cognitive ability, communication, use of language, skills, talents, and behaviour vary (Ousley & Cermak, 2014). The heterogeneity of autism contributed to the changes from categorisation to the dimensional classification of autism (Ecker et al., 2010; Lord et al., 2012) in the ICD-11 and DSM-5. By recognising that autism is a continuum, other subgroups (Asperger's syndrome, autistic disorder, and pervasive developmental disorder), previously separately delineated in the ICD-10 and DSM-IV, were consolidated into the DSM-5 and ICD-11 as an autism spectrum disorder. However, the ICD-11 retains categories such as 6A02.0 (autism spectrum disorder without disorder of intellectual development and with mild or no impairment of functional language). The extent of variability determines the level of support indicated in the DSM-5 as the degree of severity from level 1 to 3; *support, substantial support or very substantial support* (APA, 2013, p. 52).

1.3 Autism Prevalence and Incidence Rate

The WHO states that about 16% or 1.3 billion people globally have some form of disability (Barrett, Kamenov, Pearce & Cieza, 2022). Of these, they estimate a 1:100 rate for autism.

There is no unified data on prevalence and incidence rates for autism, perhaps due to study methodologies. The rates are dependent on the origin of the study, the methodology of identifying the sample, the measures of autism employed, and the cut-offs used. In Taiwan, a population-based study with 372,642 participants aged 18 years and below reported an increased rate of 1.79 to 28.72 per 10 000 from 1996 to 2005 (Chien, Lin, Chou & Chou, 2011). In the UK, the rates were – 1: 100 (Baird et al., 2006), and the 2010 USA rates were 1: 68 (Baio, 2014). Russell et al. (2022) reported an astronomical increase of 787% in the apparent incidence of autism diagnoses between 1998 and 2018 in the UK. For the US, the

rate as of 2018 for children aged eight years is now estimated as 1: 44 (Maenner et al., 2018). In Denmark, a recent register-based cohort study with 6,989,627 residents showed a prevalence rate of 400 per 100,000 for autism (Weye et al., 2021). In Sweden, the 2011 rates were found to be between 0.40% and 2.46% among 0 –27-year-olds (Idring et al., 2015). Solmi et al. (2022) has more on prevalence, incidence, and global burden rates for over 200 countries. These variabilities in prevalence rates may be attributed to various factors, such as increased awareness and diagnosis (Idring et al., 2015), increased interest in the presentation and diagnosis of autism in females, who might have been previously left out in epidemiological and clinical studies (Solmi et al., 2022), age of participants, study location and diagnostic criteria, different measures used by the different researchers, the interpretations given to the diagnostic criteria or the definitions of ASD amongst other factors (Williams, Higgins, & Brayne, 2006).

Unfortunately, there is still a brevity of data on both incidence and prevalence rates in Sub-Sahara Africa, including Nigeria, the focus of this thesis (Abubakar, Ssewanyana, de Vries & Newton, 2016a; Bakare & Munir, 2011). However, Lotter (1978) indicated that about 0.7% of children with intellectual disabilities might have ASD. Some of the limitations for the unavailability of data are the lack of screening and diagnostic tools; moreover, where there are tools, there is a lack of adequately trained professionals to administer them. Although Nigeria's prevalence rates are presently unknown, the number of children affected will likely reflect global trends. However, the global rates in and of themselves may not be entirely accurate as data from Africa is missing (Abubakar et al., 2016a).

1.4 Autism and Common Comorbidities

In addition to the core features described in the ICD-11 and DSM-5, there are other developmental disorders, challenges or higher language and cognitive skills that can coexist with autism in an individual (Ousley & Cermak, 2014; Lord et al., 2012). These developmental disorders and behaviours include but are not limited to anxiety, attention deficit hyperactivity disorder (ADHD), self-injurious behaviours, sleep issues, eating disorders, physiological and medical conditions (WHO, 2020, 6A02). Furthermore, the ICD-11 and DSM-5 allow for the specification of autism with or without accompanying structural and functional language impairment and with or without accompanying intellectual disability (ICD-11, 2020, 6A02.0, 6A02.1, 6A02.2, 6A02.3, 6A02.5; APA, 2013, p.51).

1.5 Theories of Autism

Autism spectrum disorder is a complex condition requiring dedicated research. To unravel the complexity and better understand the disorder, scholars and experts have formulated various theories about autism (Vivanti & Messinger, 2021). Since Leo Kanner and Hans Asperger first described autism, many theories have been propounded. These theories have gone from the intriguing and inane to the advanced, from 'refrigerator mom' to genetic factors. There are about twelve theories which fall under three major categories – biological, environmental, and neuropsychological. A few significant theories are described considering their relevance for screening for autism.

1.5.1 Biological and Environmental Theories

Out of the three strands of biological theories – structural abnormality, functional/chemical abnormality and genetic differences, there is a body of research focused on genetic differences (Ha, Sohn, Kim, Sim & Cheon, 2015; Marotta et al., 2020; Vivanti & Messinger, 2021). Rutter, Bailey, Bolton & Le Couteur (1994, p. 311) stated that "the organic basis of autism is no longer in dispute and is a matter of common consent". Over the years, various studies, including twin studies, have identified the genetic factors in autism (Folstein & Rutter, 1977; Bailey et al., 1995; Geschwind, 2011). However, unlike disorders like sickle cell anaemia, Down syndrome and other known genetic conditions, the vast number of biomarkers associated with the disorder makes it challenging to attribute autism to a single gene (Pasco, 2010; Fu, Satterstrom, Peng, et al., 2022). Others have proposed congenital causes such as increased foetal testosterone (Baron-Cohen et al., 2020) and, most recently, connections between autism and assisted pregnancies – IVF (in vitro fertilisation) and ICSI (intracytoplasmic sperm injection) were hypothesised. Studies to debunk the hypotheses exist (Djuwantono, Aviani, Permadi, Achmad & Halim, 2020; Sandin, Nygren, Iliadou, Hultman & Reichenberg, 2013). "Compared with spontaneous conception, IVF treatment overall was not associated with autistic disorder but was associated with a small but statistically significantly increased risk of mental retardation" (Sandin et al., 2013, p. 75). As a spectrum condition, autism affects everyone differently and impacts skills, abilities, and behaviours in many ways. Autism is a neurodevelopmental disorder with complex genetic aetiology. The biological link which has become accepted continues to be researched.

Some work has been done regarding structural abnormality as a possible cause of autism. The earlier studies focused on cerebellar abnormalities – hyperplasia (overdevelopment) or hypoplasia (underdevelopment) of the brain. For instance, Gaffney, Tsai, Kuperman, & Minchin (1987) reported a difference in the size of the cerebella between those subjects with and without autism – the cerebella of the autistic persons were proportionally smaller, and the fourth ventricles proportionally larger. Courchesne, Townsend, & Saitoh (1994) reported the results of 78 autistic persons from four separate studies, which revealed the presence of hypoplasia in 84% to 92% of the patients and hyperplasia in 8% to 16% of the patients. However, a similar study conducted around the same timeframe (Rapin & Katzman, 1998), while acknowledging the possibility of maldevelopment of the cerebellum and limbic structures, suggested further research as too few brains had been examined and without modern equipment. More recent studies (Hazlett et al., 2017; Zheng, Eilam-Stock, Wu, Spagna, Chen, Hu & Fan; 2019) have also examined this notion of head and brain size as a possible cause of autism, and while some connection appears to exist, the exact timing of the development of behavioural symptoms of autism remain unclear. Closely related to the structural abnormality theory is the chemical/functional abnormality theory. A suggestion of this theory is that genetic mutation in early brain wiring, and synapse formation increases the risk of developing autism (Robichaux & Cowan, 2013). But Akshoomoff, Pierce & Courchesne (2002, p.628) concluded that "anatomical abnormalities in autism result from the developmental dysregulation of brain growth, the causes of which are genetic defects that produce detectable abnormalities in molecular brain growth factors". Some studies, however, suggest the possible connections between biochemistry, brain anatomy and physiology, and genetics in the diagnosis of autism (Belmonte et al., 2004).

Environmental theories have postulated that air pollutants, parental age, toxins, or chemicals may cause autism (Hertz-Picciotto, Schmidt & Krakowiak, 2018). Other possible causes suggested include drug interactions during pregnancy, vaccines, maternal stress, mineral deficiencies, heavy metals, and exposure to infectious diseases (Rodier & Hyman, 1998; Grabrucker, 2013). Of interest are the studies conducted in Nigeria (Blaurock-Busch & Nwokolo, 2018) and Japan (Yasuda, Yoshida, Yasuda & Tsutsui, 2011) using plasma zinc samples from participants, which suggested that infantile zinc deficiency may epigenetically contribute to the pathogenesis of autism. However, no significant relationship between zinc and autism has been found especially given the sampling method (Babaknejad, Sayehmiri, Sayehmiri, Mohamadkhani & Bahrami, 2016; Abdulla, 1983). Although environmental risk

factors have been suggested based on human and animal research, many more, and perhaps the most significant risk factors, remain unidentified. While a combination of environmental and genetic factors may contribute to the development of autism in some individuals, results remain inconclusive as more studies are required (Bai et al., 2019; Dietert, Dietert & Dewitt, 2011). Thus, as interesting and plausible as the environmental theory sounds, there remains insufficient evidence for its complete support (Rutter, 2011). "While specific genetic mutations, CNVs, and environmental exposures appear to be strong causal factors, they are only in rare cases" (LaSalle, Vallero & Mitchell, 2013, p. 106).

1.5.2 Neuropsychological Theories

Several neuropsychological theories have been advanced to explain the impairments in autism (Vivanti & Messinger, 2021; Rozga, Anderson & Robins, 2011). Initially, these theories were based on the triad of impairments (Rao, Mysore & Raman, 2016; Happé, 1994) but have been expanded to include cognitive and social deficit theories. Included are theories of executive function, information processing, social motivation, social cognition, systemising-empathising, context blindness, monotropism, and intense world theory. Neuropsychological theories can be categorised into two main groups; those based on cognitive disturbance and those emphasising early social deficits. Cognitive disturbance theory covers executive function and information processing, while early social deficits deal with social cognition and social motivation. These four, along with other dominant theories from the research, will be discussed.

1.5.2.1 Executive Function Theory

The executive function theory (EF) suggests that the deficits in ASD reflect the impairment of higher-order cognitive skills needed to plan and generate goal-directed behaviour (Lezak, Howieson, Loring, Hannay & Fischer, 2004). These deficits are reflected in areas related to working memory, planning, inhibition, cognitive flexibility, and self-monitoring. The theory also suggests that many of the restricted interests, stereotypic or repetitive behaviours that are often features of autism, are like the behaviours seen in people diagnosed with executive dysfunction. Persons diagnosed with executive dysfunction are likely to have problems with impulsive behaviour, personality changes, reduced creativity, planning and flexibility, and difficulties with change (Ozonoff, Roger and Pennington, 1991). However, executive dysfunction is characteristic of individuals who have suffered damage to the prefrontal cortex; whilst in autism, it has been proposed that the individual's prefrontal cortex may not have matured in a typical way (Ha et al., 2015; Ecker et al., 2010; Mesulam, 2002). Ozonoff

et al. (1991) define executive function as "the ability to maintain an appropriate problem-solving set for attainment of a future goal; it includes behaviors such as planning, impulse control, inhibition of prepotent but irrelevant responses, set maintenance, organised search, and flexibility of thought and action" (p.1083). Further examination of this theory has seen individuals tested on verbal fluency, problem-solving ability, working memory, response inhibition and set-shifting (Steel, Gorman & Flexman, 1984; Demetriou, DeMayo & Guastella, 2019; Rozga, Anderson & Robins, 2011). While there is no clear consensus on the outcomes, findings indicate an agreement on the importance and contribution of EF in studying autism (Hill, 2004; Joseph & Tager-Flusberg, 2004).

Executive function theory may help explain some deficits in persons with autism, but the theory cannot explain the deficits in all domains due to the heterogeneity of the spectrum. For instance, some studies have shown that EF can explain deficits in repetitive and stereotypic behaviour but not social domains (Lopez, Lincoln, Ozonoff, & Lai, 2005; Rozga et al., 2011). There also remains a gap in explaining the aetiology and symptomatology of autism based on EF theory. Studies have typically focused on autism symptoms to the exclusion of comorbidities in participants (Lopez et al., 2005). The executive function theory does not reliably help with understanding the full range and interaction of features typically seen in individuals with autism (Rozga et al., 2011; Joseph, 1999) and deficits in executive function may present differently in autistic individuals than in persons with other developmental disabilities (Hill, 2004; Pennington & Ozonoff, 1996; Rozga et al., 2011). While the EF theory has provided some insight into deficits seen in autism, more research around the theory is required to fully understand its role in the symptomatology of autism (Joseph, 1999).

1.5.2.2 Information Processing Theory

While the executive function theory has been used to explain the difficulties that autistic people may have with cognitive flexibility and planning, it does not explain why they might also have relative strengths, particularly around visuospatial awareness, nor does it explain the non-social features of autism.

Another feature of autism is the ability to focus on specifics (local processing) rather than the whole picture (global processing); this includes the inability to integrate pieces of information into a meaningful whole. Two models have been used to describe this, complex information processing and central coherence, both incorporating aspects of the executive functioning theory. The complex information processing model suggests that autism is a

disorder of higher-order (cognitive and neurological) integrative functioning (Williams, Goldstein, & Minshew, 2006; Minshew, Johnson, & Luna, 2001). The inability to process complex information is usually reflected in some people during problem-solving, concept formation or complex memory, a life-long challenge.

On the other hand, the weak central coherence (WCC) theory hypothesises that people with autism struggle to integrate pieces of information into a meaningful whole (Frith, 1989). Research has tended to support the WCC theory (Vivanti & Messinger, 2021; Bölte, Holtmann, Poustka, Scheurich, & Schmidt, 2007; Hill & Frith, 2003; Happé & Frith, 2006), submitting that autistic individuals are highly skilled at tasks requiring attention to detail and less skilled on those requiring attention to the whole. The inability to see the whole may explain the difficulties people with autism have with understanding the subtleties of communication and language.

The complex information-processing and WCC theories can account for some patterns of strengths and weaknesses observed in autism. WCC theory does this by proposing different abilities in processing information – global or featural – while complex information processing theory highlights differences in processing information that centre on levels of complexity. A fuller and more robust account of the informational processing biases that may underline the behavioural manifestations of autism may be reached if both complex and WCC theories are combined (Rozga et al., 2011).

1.5.2.3 Predictive Coding Theory

A follow-up theory to the information processing theory, the predictive coding theory (PCT), also referred to as the Bayesian Brain theory, suggests that "an autistic person's brain does not form accurate predictions or that sensory input overrides … internal predictive models" (Musser, 2019, p1). Others (Loannou et al., 2020; Van de Cruys et al., 2014; Cannon, O'Brien, Bungert & Sinha, 2021; Sinha et al., 2014) have argued and hypothesised that PCT explains some of the difficulties experienced by persons with autism in processing previously acquired information and predicting events with precision while juxtaposing the outcomes of their behaviours with expected social responses (Lawson, Rees & Friston, 2014; Ganglmayer, Schuwerk, Sodian, & Paulus, 2020). Additionally, PCT challenges the idea that persons with autism have specific neurocognitive impairments (Constant, Bervoets, Hens, & Van de Cruys, 2020; Karvelis, Seitz, Lawrie, & Seriès, 2018) but rather suggests that the weight or 'predictive precision' allocated to each sensory input is relative. Because persons with autism often tend to process information as almost entirely

new, the precision with which they do so is suboptimal relative to neurotypical individuals (Loannou et al., 2020).

By focusing on predictive accuracy, proponents of PCT submit that delayed responses to sensory and neural inputs, can potentially explain the social communication and interaction impairments and restricted and repetitive interests of person with autism (Ganglmayer et al., 2020; Sinha et al., 2014; Van de Cruys et al., 2014). Also, the reduced predictive accuracy in persons with autism should not be seen as a deficit but rather a different application of previous information (Ganglmayer et al., 2020). Different studies examining the PCT appear to have arrived at similar conclusions - the possibility of improved understanding of autism and a shift in practice (Cannon et al., 2021; Van de Cruys et al., 2014; Loannou et al., 2020; Constant et al., 2020).

1.5.2.4 Social Motivation Theory

The social motivation theory posits that individuals with autism find social stimuli less rewarding than neurotypical people. Another posit of the theory is the concept that most people are typically motivated to orient towards social stimuli (faces, voices, cues, and gestures) and that this is interrupted or otherwise altered in some way in people with autism by some biological mechanism (Clements, Zoltowski, Yankowitz, Yerys, Schultz, & Herrington, 2018; Chevallier, Kohls, Troiani, Brodkin & Schultz, 2012). Additionally, proponents of the theory argue that the lack of early emergence of social engagement in autism limits the child's social experiences, thus, inhibiting the development of neural pathways which form because of social input (Clements et al., 2018; Mundy, 1995). Other studies have examined the values of social stimuli and engagements, such as seeking, liking, and valuing the rewards (Chevallier et al., 2012; Abrams et al., 2013), concluding that persons with autism who have deficits in emotional systems lack appropriate social communication skills. Social Motivation theories focus on a biologically based, early emerging lack of motivation for social engagement in autism. This deficit interrupts the development of neurodevelopmental pathways that underlie a wide range of social-communicative behaviours known to be diminished in people with autism. Importantly, social motivation theories aim to explain both the presence and the emergence of behavioural and neural manifestations of autism and how they are related to one another.

While social motivation theories seem to account satisfactorily for social-communicative impairments observed in autism, they do not adequately explain repetitive behaviours and restricted repertoires or cognitive profiles. It also appears that social motivation theories do

not address the heterogeneity of autism and the interplay between skill acquisition and the environment (Livingston, Shah, & Happé, 2019). Most importantly, insufficient and inconsistent evidence challenges the support of the theories (Bottini, 2018; Clements et al., 2018). Methodological limitations are a possible factor for the inconsistency in findings, but these remain unclear.

1.5.2.5 Social Cognitive Theory

The focus of social cognitive theories is the concern with impairments in an individual's ability to represent and think about the thoughts, beliefs and feelings of others, a key to managing and engaging in everyday social interactions. Some individuals with autism have been shown to have deficits in intuitive deductions, imaginative responses, pretend play, non-literal language, understanding sarcasm, and understanding of irony (Hill & Frith, 2003; Leslie & Happé, 1989; Happé, 1995). These impairments are observed in social contexts where individuals with autism cannot regulate social behaviours. This deficit is referred to widely in the literature as Theory of Mind (ToM) and sometimes as mind blindness. Premack & Woodruff (1978) defined ToM as an individual's ability to "impute mental states to himself and others" (p. 515). Baron-Cohen, Leslie & Frith (1985) further described the failure of autistic children to assign thoughts and beliefs to others as an "inability to represent mental states" (p. 43). Generally, this deficit disadvantages autistic individuals who experience difficulties socialising and understanding and predicting other people's behaviours. Early research proposed that a specific cognitive mechanism underpinned by identifiable neural structures was responsible for ToM (Baron-Cohen, Tager-Flusberg, & Cohen, 1994). Later studies suggest that impairments in social perception and behaviour in autism are linked with 'social brain' systems underpinning social cognition (Critchley et al., 2000). Studies on ToM have shown that the developmental timeline is from early infancy to age five years; however, in individuals with autism, it takes longer and lasts throughout their lifetime (Astington, & Edward, 2010; Slaughter, 2015; Moran, 2013). Although some more able individuals may acquire 'a conscious theory of mind', they still lack the intuitiveness required in daily life (Hill & Frith, 2003, p. 3). As a result of various studies, ToM offers an acceptable explanation for the difficulties individuals with autism experience in understanding the beliefs, feelings, and emotions of others. However, other domains of autism cannot be explained by ToM.

1.5.3 Empathising – Systemising Theory

The empathising–systemising (E-S) theory proposes that people may be classified along two dimensions: empathising and systemising. It measures an individual's strength of interest in empathy (the ability to identify and understand the thoughts and feelings of others and respond with appropriate emotions) and an individual's strength of interest in systems (the drive to analyse or construct them). Baron-Cohen (2009) hypothesised that one strength of the E-S theory is that it can clarify both the social and non-social features of autism. Baron-Cohen (2009) also posited that individuals with autism show delays and impairments in empathising and superior skills in systematising and that the extent of discrepancy between these two abilities is indicative of autism. Below-average empathy is seen as a simplified way to explain social and communication difficulties. In contrast, average or even above average systemising is used to explain the narrow interests, repetitive behaviours, and resistance to change or the need for sameness. E-S theory postulates that everything is kept constant in systemising, and only one thing changes at a time. The E-S theory also sought to explain ToM (Baron-Cohen et al., 1985) and cognitive differences between sexes (Baron-Cohen, 2002). He submitted that the cognitive challenges in autism appeared in areas in which he argued that the average female outperformed males and suggested this was why cognitive strengths in autism appeared in areas where males, on average, outperformed females (Baron-Cohen, 2002).

In summary, E-S theory can help characterise the unique profile of people with autism and explain an inability to generalise in autism. A few other disabilities include empathy difficulties, but arguably only individuals with autism demonstrate the disconnection between this and their integral systemising drive (Merritt, 2012).

1.5.4 Monotropism

The theory of Monotropism, an interest-based account of autism (Murray, 2018; Murray, Lesser & Lawson, 2005), argues that the core feature of autism is the "difference in the strategies employed in the distribution of scarce attention" (p. 140). In other words, monotropy is the restricted pattern or range of interest. They also suggested that attention, measured by focus, may be broadly distributed over many interests or concentrated in a few interests. Murray et al. (2005) further suggest that task demand drives the availability and allocation of attention for any person at any given time, as evidenced in other studies (Sergeant, 2000; Davidson, Alais, van Boxtel & Tsuchiya, 2008). Murray et al. (2005) and Milton (2012) described monotropic behaviour in autism as a different cognitive style rather

than an impairment. Monotropy, therefore, suggests that the 'restricted repetitive and stereotyped patterns of behaviour, interests and activities' in autism are "deep basins of attraction where attention gets caught and may be expressed in a thought or action over and over again" (Murray et al., 2005, p. 146) with no apparent alternative attractor. Because attention-shifting is required in daily activities such as conversations, social interactions, and information processing, an individual with autism may find these settings challenging. However, Murray et al. (2005) argued that individuals with autism could complete such tasks once they understand the goal, are motivated by it, understand precisely what task it is and what steps must be taken to carry it out and can take the identified steps (p. 141).

Another feature of autism that monotropism attempts to explain is the qualitative impairment in communication. Here Murray et al. (2005) argued that the perceptions and thought patterns in autism are a fragmented understanding of the environment, unlike in neurotypicals with flexible rather than tightly focused attention. This focused attention leads the autistic individual to miss useful information as the sequencing of events is unrecognised due to a lack of cognitive connection between sequence elements. Conversational skills require sequencing of information – tone, sound, rule, grammar, word meaning, sentence meanings, and adjusting to each other's current interests, which can be overwhelming for autistics. Each change during conversations requires adjustment, and for persons with autism, they require time, and if the current and leading interest is not engaged strongly enough, attention is shifted (Murray et al., 2005). "Unless language becomes an object of interest it will take monotropic individuals longer to realise that language is meaningful" (Murray et al., 2005, p. 150).

Monotropism highlights the link between restricted repetitive and stereotyped patterns of behaviour, interests and activities in autism and provides insight into difficulties with change and attention-shifting. Monotropism also suggests that the "acceptance of uncertainty and unpredictability and the existence of categorical uncertainty need to be taught" (Murray et al., 2005, p. 152), thus, allowing individuals with autism to acquire relevant skills. The theory of monotropism provides evidence for the uneven cognitive profile in autism as a hyperfocus on areas of interest leading to higher abilities. The theory also suggests several features found in autistic personal accounts not addressed by the other theories.

1.5.5 Relationships between the Theories

Understanding autism has involved an enduring and longstanding discourse with different aetiological theories postulated. These theories stem from different perspectives and contexts upon which research and practice in autism are based. The DSM-5 core symptoms of autism include social communication deficits and restricted repetitive and stereotypic patterns of behaviours, interests, and activities. However, the features and causes of autism are not embodied in one theory, but sometimes there are overlaps.

ToM explains the pragmatic impairments of language and communication in terms of social deficits starting from the idea that mental states must be inferred, thus, requiring a complex cognitive mechanism (Baron-Cohen et al., 1985). While the WCC focuses on global versus local processing, it also submits that the ability to process complex information varies in individuals depending on complexity (Williams et al., 2006). Monotropism, on the other hand, states that processing ability depends on the availability of attention (Murray et al., 2005). Amongst other difficulties, EF postulates the reduced flexibility in persons with autism and their difficulties with changing positions. Monotropism describes this inflexibility as an intense focus driven by competition between mental processes; thus, resistance to change should not be viewed as a deficit. Comparing WCC with E-S, Baron-Cohen (2002) argues that non-social cognitive impairments, which present as local processing, are the starting point of systemising for persons with autism. It, therefore, seems that WCC provides some foundation for E-S in this regard. While all other theories mentioned have focused on social communication, language and repetitive behaviours, E-S theory, an extension of ToM (Baron-Cohen, 2009), also suggested the inability of individuals with autism to generalise. Monotropism describes this inability to generalise as a "corollary" of the tight focus (Murray et al., 2005). The WCC and EF are similar in that they postulate broad, domain-general cognitive impairments in information processing or, in the case of EF, in executive control over information processing and the planning of actions.

Other studies (Vivanti & Messinger, 2021; Joseph, 1999) have compared the biological with the neuropsychological theories, concluding that while neuropsychological theories have succeeded in describing the features of autism, they have failed to explain other domains, such as repetitive behaviours. They submit that repetitive responses are best explained by unconnected neural pathways and developmental failures (Levy, 2007).

Early studies attempted to draw similarities and contrast between the major theories – EF, WCC and ToM but submit that they remain distinct in their assertions (Rajendran & Mitchell, 2007). Belmonte et al. (2004) noted that regardless of which theory, the field of autism suffers from a lack of unification of the different theoretical concepts, including EF, WCC, ToM and E-S. A weakness relevant to all the theories reviewed is the lack of data connecting the neuropsychological deficits that have been identified in autism to the behavioural abnormalities they presumably underlie.

1.6 Intellectual Disabilities

Historically, individuals with an intellectual developmental disorder/intellectual disability have been referred to in different ways, even in legislation, such as subnormal, mentally handicapped, mentally retarded, idiot, moron, and imbecile. These pejorative labels had undertones of intolerance and unkindness from society, causing the individuals to be viewed as a burden to their families and society. Philosophers of Roman and Greek descent whom highly valued reasoning abilities ridiculed people with intellectual disabilities (Mirabi, 1985). Similarly, in England, such persons were seen as 'fools' or 'mad', and these two conditions were not always distinguished, even though as early as the 13th century, documents distinguishing between people with intellectual disabilities ('natural fools') and people with mental illness ('lunatics'), had appeared. 'Fools' were thought incapable of rational behaviour, and if they could not live with their families, they were incarcerated in some periods (Narby, 1982). The United States was not left out, as such persons were regarded as 'feeble-minded', and eugenic attitudes flourished (Robert & Kurtz, 1987). As the West departs from its previous constructs, Africa gradually follows suit, although at a slower pace. At times, the concept of 'madness' is still attributed to such persons in Africa.

In the late 1600s, a neuroanatomist Thomas Willis gave some of the first insights into the aetiology of intellectual disability as a disease in his work *Pathologicae cerebri, et nervosi generis specimen*, on the pathology and neurophysiology of the brain (Molnár, 2004). Later, in 1866, Langdon Down identified another possible cause of intellectual disability, later known as Down syndrome, in a group of individuals in Normansfield with similar characteristics to each other, and as time went on, other genetic causes were identified.

As the concept and possibility of intellectual disability as a disease gained ground, there began to be a recognition that environmental causes were also relevant, and with these came the possibility of intervention. A pioneer in this was Alfred Binet in France, who, in the

early 1900s, was asked to consider how to identify children who needed special schooling after France introduced universal education, and he went on to develop the first IQ test. However, his findings and methods were to herald a period of institutionalisation of people with intellectual disabilities and widespread eugenic attitudes in the USA and elsewhere, and it was not until the normalisation and civil rights movements of the 1960s that there was an emergence of new terminology and institutional changes, alongside the emergence of state-funded community care. Journals and institutions gradually changed their names (Schalock, Luckasson & Shogren, 2007; Nash, Hawkins, Kawchuk & Shea, 2012). For instance, the Australian Journal of Mental Retardation (1970 – 1979) became the Journal of Intellectual & Developmental Disability (1996 – current), the British Institute of Mental Handicap now the British Institute of Learning Disability (1982 – 1993) and the American Association on Intellectual and Developmental Disabilities (2007 – current) which previously had a variation of title that included "idiotic, feebleminded Persons and mental retardation". World Health Organization's ICD-10 and American Psychiatric Association's DSM-IV used "mental retardation"; however, in the revised ICD-11 and DSM-5, the terms "disorders of intellectual development" and "intellectual disability (intellectual developmental disorder)" are used, respectively.

In the United States (US), the broader term developmental disability covers a wide range of disorders, including autism, intellectual disability, cerebral palsy, and other disorders occurring in the developmental window of birth to age 18 years. However, in school settings, intellectual disability is used synonymously with persons whose cognitive abilities are two standard deviations below the mean. In the United Kingdom (UK), learning disabilities is the official term for intellectual disability, while learning difficulties refer to more specific cognitive disorders, such as dyslexia. Intellectual disability, the preferred terminology among international researchers and advocates, will be used in this research.

1.6.1 Characteristics of Intellectual Disability

Under the DSM-5 intellectual disability has no specific age of onset, but impairment in adaptive behaviour and general mental functioning should be present before the age of 18 years. Impairment in adaptive behaviour is compared to the individual's peers by age, gender, and socio-culturally. Features in the following domains are listed (APA, 2013, pp.37-38).

1. The conceptual domain, which covers skills in language, reading, writing, math, reasoning, knowledge, and memory.

2. The social domain, which covers interpersonal communication, empathy, making and retaining friends, and social judgment.
3. Adaptive skills, which cover self-help skills, self-management, organisational skills, community skills, money management and similar behaviours.

While intellectual disability does not have a specific age requirement, an individual's symptoms must begin during the developmental period and are diagnosed based on the severity of deficits in cognitive skills and adaptive functioning. Intellectual disability can co-occur with other conditions, such as autism, anxiety, depression, and hyperactivity.

Within the ICD-11, intellectual disability is described as a "group of etiologically diverse conditions originating during the developmental period" (6A00), excluding dementia. Individuals with intellectual disabilities should have general intellectual functioning and adaptive skills below the average range, where the bottom of the average range is set at two or more standard deviations below the mean. The ICD-11 and DSM-5 specify the extent of impairment as either mild, moderate, severe, or profound based on the combination of general intellectual and adaptive functioning and no longer by intelligence quotient (IQ) scores. The level of support given to the individual is determined by adaptive functioning, particularly as IQ scores are deemed problematic and not sensitive enough at the very low ability range (Hessl et al., 2009).

1.6.2 Adaptive Behaviour

Adaptive functioning (DSM-5) and adaptive behaviour (ICD-11) will be used interchangeably within this study as both are used in research.

Adaptive behaviour is defined as the "extent to which a person is capable of being self-sufficient in real-life situations, including the functional use of communication, socialisation, daily living and motor skills" (Kanne et al., 2011). It is also defined as a "collection of conceptual, social, and practical skills that have been learned and are performed by people in their everyday lives" (Tassé et al., 2012, p. 291). The authors of the Vineland Adaptive Behavior Scales, 3rd Edition (VABS-3) defined adaptive behaviour as "the performance of daily activities required for personal and social sufficiency" (Sparrow, Cicchetti & Saulnier, 2016, p. 10).

The construct of adaptive behaviour and its measurement has been revised over several years (Tassé et al., 2012). Sparrow et al. (2016) summarised the reviews under four

principles – firstly, it is age-related, and the complexity of skills increases with age. Secondly, adaptive functioning is evaluated in a social context; thirdly, adaptive behaviour is modifiable; lastly, it is defined by typical performance, not ability. In young children, self-care skills such as dressing and using the toilet are relevant, but in adults, behaviours such as money handling or job retention become relevant. In the social context, expected behaviours are defined by others and are therefore highly culturally dependent. For instance, offering an older adult a seat as required or running errands for the elderly is important in some cultures. Where adaptive functioning is low, skills can often be taught through intervention, but they can also be eroded because of neglect or trauma. The ability to perform a required task is necessary but not sufficient in the sense that some individuals may have limitations imposed by others (a bevvy of domestic staff, skills not taught, a custodial environment where tasks are done by others) which impedes performance. Persons with intellectual disabilities or autism, who are not given opportunities or taught adaptive skills, will score low when assessed.

As previously mentioned, although evidence of a low IQ is still required for a diagnosis of intellectual disability, both the ICD-11 and DSM-5 stipulate that the level of intellectual disability be based on adaptive functioning. Thus, reflecting the understanding that although cognitive and adaptive functioning are correlated, skill acquisition capacity may differ from its execution in everyday life (Sparrow et al., 2016; Alexander & Reynolds, 2020). Studies have also shown that adaptive behaviour can be significantly impaired even among persons with autism and other neurodevelopmental disabilities with high IQ but no intellectual disability (Meyer, Powell, Butera, Klinger, & Klinger, 2018; Duncan, Ruble, Meinzen-Derr, Thomas, & Stark, 2018). Therefore, adaptive functioning, not IQ, is most associated with functional outcomes (Zheng, LeWinn, Ceja, Hanna-Attisha, O'Connell, & Bishop, 2021). An individual's adaptive functioning is determined based on observation, adaptive behaviour assessments, skill assessments and description of abilities provided by someone familiar with the individual.

1.6.3 Intellectual Functioning

The APA Dictionary of Psychology defines intellectual functioning as "any of the mental functions involved in acquiring, developing, and relating ideas, concepts, and hypotheses". Intellectual functioning, also referred to as intelligence, is a general mental capacity that includes planning, problem-solving, reasoning, abstract thinking, and comprehending

complex ideas (Carr, Linehan, O'Reilly, Walsh, McEvoy, 2016). It can be determined by individually administered general intelligence tests such as IQ tests. The IQ is a score determined by an individual's performance on a standardised intelligence test relative to the average performance of peers of the same age. Various IQ tests, such as the Weschler Intelligence Scales, Woodcock-Johnson Test of Cognitive Abilities and the Kaufman Tests, are available with varying IQ ranges. However, where tests use an average score of 100 and a standard deviation of 15, a score of 70 and below indicates a significant impairment of general intellectual functioning.

1.7 Intellectual Disability Prevalence and Incidence Rate

As previously mentioned, WHO states that about 16% or 1.3 billion people globally have some form of disability (Barrett, Kamenov, Pearce & Cieza, 2022). Of these, they estimate that 2% – 4% have functioning difficulties.

Like autism, data on the prevalence of intellectual disability in different studies are challenging to compare, as the definition of intellectual disability drives study methodologies, diagnostic criteria, diagnostic tools, the conceptualisation of intellectual disability, and the identification of those with intellectual disability (Bertelli et al., 2022). Maulik et al. (2022) provides rates for some developed economies – Finland & Netherlands currently have a prevalence of less than 1%. In the USA, studies revealed a prevalence rate of 7/1000 among children 3–17-year-olds from 1996 to 2008 and an estimated 6.7/1000 from 2006 to 2008.

A population-based meta-analysis covering 1980 to 2009 (Maulik et al., 2011) reported a global rate of 10.37/1000, with the majority of these individuals in low (16.41/1000) and middle-income economies ((15.94 /1000) of which Nigeria is one. However, these figures cannot be assumed accurate due to insufficient studies and data out of Africa. Another systematic review (McKenzie, Milton, Smith & Ouellette-Kuntz, 2016), this time of studies between 2010 – 2015, reported a prevalence range between 0.05% and 1.55% globally. These data are, however, based on studies from eight countries – Canada, China, India, USA, Australia, Denmark, Norway and Taiwan.

The qualitative descriptor 'borderline or very low' IQ, used in some studies, should not be taken to mean 'intellectual disability'. Individuals who have 'borderline' or 'very low' (the more recent descriptor) Iqs have Iqs that are 1 SD below the mean, rather than 2 SD below

the mean as in ‘intellectual disability’. The distribution of Iqs in a population is typically Gaussian (Grabinski, & Klinkova, 2020; Patel, Apple, Kanungo & Akkal, 2018; Vock, 2008), such that approximately 67% of the population have Iqs within 1 SD from the mean (above or below), while approximately 95% of the population have Iqs within 2 SDs from the mean (above or below). So a lot more people have Iqs that are 1 SD below the mean, than 2 SDs below the mean (the criteria for an actual intellectual disability).

Carr & O’Reilly (2016) reported that about 85% of individuals with intellectual disability are classified within the mild range (IQ of 50 - 69), 10% within the moderate range (IQ of 35 - 49), 4% within the severe range (IQ of 20 - 34) and 2% within the profound range (IQ <20). Additionally, there is a difference between the true prevalence rate and the administrative prevalence depending on the study methodologies, case definitions and study settings (McBride et al., 2021; Ouellette-Kuntz et al., 2009). True prevalence rates are typically determined through clinical screening and diagnosis while administrative rates are based on defined geographic areas and counting those receiving services or identified as eligible for services (Ouellette-Kuntz et al., 2009). For instance, McBride et al. (2021) observed a decrease in the total population prevalence rate from 2% to 0.5% when they compared the administrative data to the clinical data. Similarly, Ouellette-Kuntz et al. (2009) observed a reduction in prevalence rate from 6.0–9.0 per 1,000 in previous studies to 4.7 per 1,000 when comparing the results of their administrative prevalence rates to the true rates. Typically, administrative rates are lower by a half or more than true rates (Carr & O’Reilly, 2016).

1.8 Theories of Intelligence and Causes of Intellectual Disability

The aetiology of intellectual disability is not entirely known, but theories and causes around biological, environmental, behavioural, and psychosocial factors have been suggested (Shree & Shukla, 2016; Parmenter, 2011). Biological causes of intellectual disability which have been mentioned in the literature include genetic mutation, chromosomal disorders, pre and postnatal conditions (Gilissen et al., 2014), while psychosocial factors include abuse, neglect, poverty, limited stimulation, and poor parent-child interactions (Shree & Shukla, 2016).

Intellectual disabilities resulting from genetic factors constituted of syndromic or biomedical and non-syndromic contributors and have been suggested to account for 20 – 50% of intellectual disability cases (Maulik et al., 2011; Huang, Zhu, Qu & Mu, 2016).

Biomedical aetiological factors of intellectual disability include Fragile X, Down's syndrome, Turner syndrome, Cri du Chat syndrome, Williams's syndrome, Prader–Willi syndrome and a host of other genomic syndromes. Although the extent of the contribution of non-syndromic factors are not fully known, so far, cognitive impairment is seen as the main factor for intellectual disability (Maulik et al., 2011; Huang et al., 2016). Other biomedical factors, some of which can be treated have been documented and these include increasing maternal age, gestational disorders (low birth weight), metabolic disorders, maternal infections, neonatal infections, birth trauma, congenital hypothyroidism, phenylketonuria, and much more (Shree & Shukla, 2016; Carr & O'Reilly, 2016; Hunag et al., 2016).One other emerging biological theory is that of DNA methylation, which suggests that a segment of the DNAs activity can be changed without impacting the sequence by the release of methyl into the body (Maulik et al., 2011; Iwase et al., 2017). Thus far, epigenetic mechanisms have been suggested to impact cognitive processes such as learning and memory (Iwase et al., 2017).

Some of the behavioural factors posited for intellectual disability includes foetal alcohol syndrome, drug use and substance abuse during pregnancy, child abuse, lack of sensory input and intellectual stimulation (Carr & O'Reilly, 2016; Emerson, 2013; Huang et al., 2016). As these behaviours are external to the person who may likely become intellectually disabled, it is possible offer some mediation and intervention as preventive measures. Another hypothesised aetiology is the relationship between social factors and intellectual disability. Social factors include poverty and its direct impacts, malnutrition, lack of access to a viable support system, stigmatisation and discrimination, inconsistent parenting due to lack of resources or neglect, and exposure to different life events which invariably stagnate the total development of an individual (Emerson, 2007, 2012 & 2013; Carr & O'Reilly, 2016). Additionally, environmental factors posited as causes of intellectual disability include exposure to harmful substances such as lead, radiation and other toxins which could damage the brain or lead to seizures (Maulik et al., 2011; Shree & Shukla, 2016). Further environmental factors also include the lack of access to adequate information and education on intellectual disability (Carr & O'Reilly, 2016).

Regardless of which theory supports the aetiology of intellectual disability, it is a challenge to effectively diagnose the condition. Several factors contribute to this, such as measurement error, evaluating the contribution of IQ scores to diagnosis, extreme scores with uneven profiles, the Flynn Effect, and test selection. In addition, determining cut-off

scores may be challenging, especially where an individual's functioning (FSIQ scores and adaptive skills scores) meets the upper limits of borderline or very low IQ or possibly slightly above these limits (Patel et al., 2018). Nonetheless, the availability of screening tools for the identification of the condition especially will help with the provision of the relevant support, education, and intervention for the Nigerian adolescent.

1.9 Comorbidity of Intellectual Disability with Autism

The ICD-11 and DSM-5 allow for specifiers of autism with or without intellectual disability and vice-versa. The ICD-11 acknowledges that individuals with autism can "present with the significant limitations in intellectual functioning and adaptive behaviour observed in disorders of intellectual development, autism spectrum disorder can also present without general limitations in intellectual functioning" (6A00). While the DSM-5 provides for the frequent comorbidity of autism and intellectual disability, it cautions about the complexity of assessing for intellectual disability due to the social communication impairments in persons with autism. Similarly, while the ICD-11 recognises the comorbidity and challenges inherent in diagnosing autism in persons with severe intellectual disability, it calls for in-depth and longitudinal assessments with great emphasis on adaptive functioning than on social skills.

Various studies have attempted to establish the connections between autism and intellectual disability and to unravel the extent of overlap (Matson & Shoemaker, 2009; Bartak & Rutter, 1976). Another factor is the heterogeneity of both disorders leading to a difference in the presentation of symptoms in individuals. This heterogeneity also makes determining comorbid conditions in older children and adolescents complex, as demonstrated in the literature (de Bildt, Sytema, Kraijer & Minderaa, 2005; Bakken et al., 2010). Some studies show that about 50 – 70% of persons with autism have an intellectual disability (Matson & Shoemaker, 2009). Other studies have also reported the proportions of people with intellectual disabilities who have autism – Bryson, Bradley, Thompson & Wainwright (2008) reported that 28% of the participants with intellectual disabilities had autism and La Malfa, Lassi, Bertelli, Salvini & Placidi (2004) reported 70%. Different factors, such as a change in diagnostic criteria, population sample, assessment methods, and understanding of symptomatology, may account for the disparities in prevalence rates (Matson & Shoemaker, 2009; Matson & Goldin, 2013).

As a result of the comorbidities, different studies focus on specific groups depending on the study aims. These groups include those with autism and intellectual disability, those with autism only and those with an intellectual disability only; however, there was no group separation in this study. As such, to better assess and tease apart these disabilities when they co-occur, various tools have been developed for assessing each condition. Tools such as the Pervasive Developmental Disorder in Mentally Retarded Persons (PDD-MRS; de Bildt et al., 2005; Kraijer & de Bildt, 2005) for screening autism in adults with intellectual disability, the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q; McKenzie & Murray, 2015) for screening intellectual disability in children and adolescents with autism. Although research is ongoing in this area and more screening tools are being developed globally, the prevalence rate for comorbidity remains uncertain. Nevertheless, continuously studying and understanding these disorders across the lifespan will increase the knowledge base.

1.10 Models of Disability

Merriam-Webster dictionary defines disability as "a physical, mental, cognitive, or developmental condition that impairs, interferes with, or limits a person's ability to engage in certain tasks or actions or participate in typical daily activities and interactions". WHO (ICF) describes disability along the dimensions of a) an impairment of an individual's body structure or function, or mental functioning, b) activity limitation, c) participation restrictions in normal daily activities, and d) environmental factors, all of which lead to the individual requiring some support. From these definitions, one can deduce that the concept of disability is fluid, and identification and intervention have been approached using different models. Disability models drive the perception, assessment, diagnosis, intervention, and general interaction with persons with any form of disability. These models are also used to provide a) definitions of disability, b) explanations of causal and responsibility attributions, and c) a guideline for the formulation and implementation of policy. There are various models, but three leading models applicable to how autism and intellectual disability are viewed in the Nigerian context are discussed.

1.10.1 Medical Model

Core ideas of the medical model of disability are that an individual's impairment can be diagnosed, cured, or rehabilitated by modern medicine or medical technology provided by sagacious professionals (Marks, 1997; Bogart, Bonnett, Logan & Kallem, 2022).

Researchers and professionals attending to individuals with autism or intellectual disability based on this model shift focus from the person to the disability, thereby viewing the disability as an illness. Regarding language, the medical model's view is perhaps the driver for some persons with autism preferring to be addressed as autistic (Rajendran & Mitchell, 2007; Milton, 2012); autism is inherently part of them and not an add-on. Medical professionals may not develop assessment measures but may focus their studies on curative measures such as drugs and genetic interventions. However, without assessments, there will be no identification of the 'disease' to be treated. Also, the position of the medical model invariably implies that disabled persons cannot actively participate in society or communities without being 'cured', thus, denying them equal access to society. In the Nigerian context, where seeking a 'cure' is rampant and sometimes ignorantly fuelled by some professionals, the environment can disable the person in a variety of ways. Environmental disablement includes factors such as the prescription of drugs rather than referring to other specialists, in some instances suggesting to families that nothing is wrong with their child, thereby dismissing their concerns, suggesting non-evidenced-based procedures or telling the families to have faith and pray. However, in the bid to find a cure for autism or intellectual disability, families may be open to assessments and diagnosis.

1.10.2 Social Model

The social model shifts the focus from the impairments of the individual to the collective responsibility of society. It posits that society's unaccommodating position in relation to individual differences causes disability (Bogart et al., 2022; Olkin, 2002). In other words, disability occurs because society assumes everyone is healthy, able, and functions according to a cultural ideology. Unlike the medical model, the social model differentiates between impairments and disabilities. While disabilities are restrictions imposed by society, impairments are seen as effects of any given condition. Therefore, the social model focuses on ultimately changing society and not curing the individual. By changing how autism and intellectual disability are viewed, society can remove the barriers to accessing the community and services, as well as identification and intervention (Rajendran & Mitchell, 2007). For Nigerian adolescents, this means a) creating more awareness in the society about autism and intellectual disability, b) educating the Nigerian society about the 'invisible/hidden' nature of autism and intellectual disability, c) providing professionals with information and relevant screening and diagnostic tools, d) engaging the government

to create relevant policies to support persons with autism or intellectual disability and e) making ‘reasonable adjustments’, for example in the workplace.

1.10.3 Religious Model

The religious model, also known as the moral model of disability, shares some similarities with the medical model in that they both allude to disability existing within the individual. The difference is that in some cultures, the religious model encourages stigmatisation. This prevalent model is the oldest worldwide (Olkin, 2002). Within this model, disability is viewed either as a punishment from God, a test of faith, the result of ancestral sins, or the sins of the disabled person (Retief & Letšosa, 2018); therefore, everything is left to God to resolve. While people’s lived experiences cannot be dismissed, often times the ‘leave it all to God’ position is misguided or stems from a misinterpretation of Biblical concepts, leading to discrimination and lack of support for persons with disability (Otieno, 2009; Creamer, 2008).

Before the advent of the Christian missionaries, various indigenous religious practices existed in Nigeria. These practices were all based on a shared belief that a supreme being is the creator. However, every person possessed a personal guiding ‘god.’ The traditional Yoruba person believed that hundreds of minor gods or spirits replaced the supreme god and influenced people’s day-to-day lives. For the Igbo traditional believers, their practice was based on the existence of numerous gods, not a single supreme being. Most of these indigenous religious practices had priestesses and priests who were purported to receive their powers by being possessed by the ‘god’ and thus had supernatural abilities. Currently, most Nigerians profess Christianity or Islam, but a few traditionalists remain. Based on 2013 estimates by the CIA, there are 51.6% of Muslims, 11.2% of Roman Catholics, 35.7% of other Christians, 0.9% of traditionalists and 0.5% unspecified worshippers (CIA, 2018). Each of these faith systems impacts the thoughts, ideas, the behavioural pattern of the people, and the government (Yesufu, 2016). Based on the religious model, whatever the disability is attributed to, the potential interference with the service persons with autism or intellectual disability require exists and cannot be overlooked.

1.11 Summary and Conclusion

In this chapter, the characteristics of persons with autism and/or intellectual disability have been described along with the theories and models that drive how they are perceived and

supported. Also, the limitations to providing prevalence data from African countries such as Nigeria are highlighted.

"The aim of theorising is usually thought to be the discovery of truth - truth about the inner relations and structure of the natural world… it is a reflective effort to see beneath the surface" (Eaton, 1921, p. 683). While a comprehensive theory of autism and model of disability continue to morph, the value of theories and models resides in their utility for guiding research and practice (van Ryn & Heaney, 1992). Theories of autism and models of disability have contributed to the evolution and development of screening tools, assessments, research, and the increasing integration of scientific knowledge and more inclusive views on intellectual disability and autism (Towle & Patrick, 2016; Brewer, Young & Lucas, 2020). However, most of the effort in relation to identification is on young children aged 12 – 36 months, with little or no focus on adolescents (Brewer et al., 2020; Thabtah & Peebles, 2019). Exploring the values of these theories and models in developing screening tools for adolescents will be useful, bearing in mind the heterogeneity and complexity of autism and intellectual disability and the relative lack of early identification in Nigeria. Chapter two focuses on the screening and diagnosis of both disorders.

Chapter 2. Screening and Diagnosis of Individuals with Autism or Intellectual Disabilities

In chapter 1, the different models of disability, theories of autism and intelligence and causes of intellectual disability were highlighted. Without identifying persons with autism or intellectual disability, an ethical, effective, and targeted intervention will be impossible.

2.1 Introduction

Detecting developmental concerns suggestive of autism or intellectual disabilities using screening tools can help identify adolescents who need further diagnosis or intervention. Screening has been widely encouraged to identify persons with autism or intellectual disability, and many screening tools are available (Thabtah & Peebles, 2019). Consensus on the ideal and practical screening tools is lacking, however, more so where the tools are used in environments other than those in which they were developed (Marlow, Servili & Tomlinson, 2019). Ideally, screening for autism and intellectual disabilities should be part of children's routine visits to health professionals at an early age, but various factors hinder this in the Nigerian context. Some of these factors include the mindset of parents or caregivers and the lack of adequate resources (Franz, Chambers, von Isenburg & de Vries, 2017.) In practice, however, adolescents are often identified when transitioning to secondary schools or facing more challenging environments and expectations of greater independence. In Africa, individuals with developmental disabilities are noticed either in schools or when parents seek medical attention for a severe illness (since visiting hospitals/health professionals for routine check-ups or minor ailments is not the norm) or when autism or intellectual disability specific research work is carried out (Knox et al., 2018; Saloojee et al., 2007; Gladstone et al., 2010; Scherzer, Chhagan, Kauchali & Susser, 2012; Bello-Mojeed, Omigbodun, Bakare, & Adewuya, 2017).

2.2 Purpose of Screening

Specific purposes for screening include a) problem identification, b) problem definition, c) referral for further investigation/diagnosis, d) inclusion for interventions, e) monitoring intervention effects, and f) evaluating outcomes (Deno, 2005). The problem identification stage is addressed by observational assessment or utilisation of a screening tool. Observational assessments rely on parent or caregiver reports or interviews about the

individual. Sometimes observational assessments are conducted by an examiner using a checklist. Observational assessments are referred to as level 1 assessments. Level 1 assessments or screening are used to probe for potential developmental disorders of various kinds. Level 2 screening is used for differentiating individuals with autism or intellectual disabilities from those who do not have autism or intellectual disability. Level 2 screening tools are also used to identify individuals who may have been missed during the level 1 assessment. Level 1 and level 2 screening tools may sometimes appear similar; however, level 2 screening tools are designed to differentiate those individuals with autism or intellectual disability from those with other developmental concerns (Brewer, Young & Lucas, 2020; Petrocchi, Levante & Lecciso, 2020). Observational assessment can be subjective and biased, considering the theories of autism and models of disability. However, it is beneficial for identifying red flags and precursors, for example, ToM (Petrocchi et al., 2020). Observational assessments are not diagnostic; they serve as pointers for further investigation. Utilising level 2 screening tools provides a more structured approach to problem identification and is less subjective. In summary, screening is a quick method for categorising individuals into groups according to some trait, characteristic or construct.

Some tools have been designed for younger children, but few to none for adolescents (Marlow et al., 2019). Those who screen positive are usually referred for a further diagnosis based on a gold standard tool. Screening is not precise but should be as accurate as possible to reduce over-referrals and under-detection. Several studies (e.g., Eldevik, Hastings, Hughes, Jahr, Eikeseth & Cross, 2009; Steiner, Goldsmith, Snow & Chawarska, 2012; Swinkels et al., 2006; Luckasson & Schalock, 2013; Schalock & Luckasson, 2013) have highlighted the benefits of early detection of developmental disabilities such as intellectual disabilities and autism. The benefits have included improved behavioural outcomes, family support, and earlier intervention. Other benefits included improved planning for educational needs and support, improved social skills, and greater cognitive and language development. Additionally, screening tools are sometimes used to monitor the progress of interventions, where the same measure is re-administered to the same individual, to examine progress, indicative of the “responsiveness” of a screening tool.

These findings have emerged predominantly from Western and high-income countries, with very limited research from low to medium-income countries (LMICs) (Tomlinson et al., 2014; Gladstone et al., 2010). LMICs are countries indexed by the published gross national income by the United Nations (United Nations, 2014; World Bank, 2020 & United Nations

Department of Economic and Social Affairs, 2021). While the presentation of autism or intellectual disability might be the same regardless of economic status, the political climate and associated social burdens within LMICs, such as in African countries, discourage the early detection of developmental disabilities as it is not seen as urgent, which increases the health disparities faced by this population (Emerson, 2012; Gladstone et al., 2014). The situation is similar for those with intellectual disabilities, with late identification leading to further intervention delay.

2.3 Methods of Screening

Screening for developmental disabilities can be done in any setting, such as the community (Kopp & Gillberg, 2011), schools (Suhail & Zafar, 2008; Webb, Morey, Thompsen, Butler, Barber & Fraser, 2003), primary care settings (Robins, 2008; Barton, Dumont-Mathieu & Fein, 2012; Gura, Champagne & Blood-Siegfried, 2011; Limbos et al., 2011), urban settings (Guevara et al., 2013), the criminal justice system (Murphy, Gardner & Freeman, 2017), and many others. Preliminary screening for intellectual disabilities or autism can occur through the use of a variety of methods, such as observation, informal and formal interviews, history taking and the use of short screening tools. Irrespective of which method is used, the essential factors to consider are the accuracy of results, validity, reliability, training requirements, ease of administration and the simplicity and ease of interpreting results (Westerlund & Sundelin, 2000; Cochrane & Holland, 1971).

As mentioned earlier, level 1 assessments are predominantly observational and can be subjective. However, using screening tools at level 2 involves using specific measures designed to be completed by the parent, a caregiver or someone who knows the individual very well. Level 2 screening tools usually include specific questions to elicit desired responses, reducing the subjectivity of responses encountered with level 1 assessments. Such measures are typically completed in 10 – 15 minutes and include the Social Communication Questionnaire, Autism Screening Quotient and Autism Spectrum Screening Questionnaire, and others (Norris & Lecavalier, 2010). The tools, as they relate to adolescents, will be discussed in Chapter 4.

Regarding intellectual disability, there is no definitive screening assessment available; as such, global developmental delay is used as a surrogate measure between 3 months and 5 years (Kishore et al., 2019). Delayed motor skills, speech, cognition, and adaptive behaviours may be assessed. Recently, screening measures administered directly to persons

suspected of having an intellectual disability have been developed. The completion time is usually between 5 to 10 minutes. Some of these measures apply to adolescents or have been developed specifically for the adolescent population and will be discussed in Chapter 4.

2.4 Screening for Autism and Intellectual Disability in Africa

Having looked at the accepted definitions of disability, intellectual functioning, the models of disability and theories of autism by the professional community and acknowledging that these are socially constructed ways of classifying people and their limitations, which have significant implications for their lives, we can understand the purpose and usefulness of screening contextually. By the early 1900s, tools for testing IQ had been developed in Western civilisations. By the 1940s, the Weschler Intelligence Scales emerged, designed for pre-schoolers, children, and adults as independent measures. Based on scores from these tests, individuals were categorised as having an intellectual disability. By the mid-'60s through the '90s, psychiatrists and psychologists also developed tools to *screen* for intellectual disability, autism, and behavioural challenges in children (Rutter, 1967; Einfeld & Tonge, 1995; Achenbach & Edelbrock, 1983; Ehlers, Gillberg & Wing, 1999).

As the years progressed, advancement in research and improved understanding of intellectual disability and autism led to the development of more assessment tools. This advancement has also made it possible to discriminate between disorders and differentiate the various levels of functioning in people with autism and intellectual disability. Not only is there a proliferation of tools, but the awareness level in Western societies led to different legal frameworks for different countries. In the US, for instance, different states mandated that children are screened during their developmental wellness visits with their paediatricians (Hacker, Penfold, Arsenault, Zhang, Murphy & Wissow, 2014; CDC, 2018). In the UK, there is NICE (National Institute for Health and Care Excellence) guidance for recognising, referring and accessing services for persons with autism. At the same time, in Canada, each province developed its own approach (Shepherd & Waddell, 2015). Regardless of the country, validated and reliable screening tools are available. Similar tools are not readily available in Nigeria and Africa as a whole.

Besides the availability of screening tools, culture and perspective as social models may contribute to individuals being screened in Western countries. Sometimes as seen in the social model of disability, the culture/beliefs of a people determine the perspective and the constructs attributed to either intellectual disability or autism. Although cultures can be

dynamic, with some beliefs changing over time, some environments are inherently conservative and tend to resist change (Inglehart & Baker, 2000). Some cultures are more resistant to change than others and promulgate laws to preserve and protect traditional cultural patterns while setting up barriers to foreign ideas (Shell-Duncan, Hernlund, Wander & Moreau, 2013; Shell-Duncan, 2008). Our cultures help us decipher the known from the unknown while framing our perspectives. People's views, ideas and beliefs about their health and disabilities are often impacted by their culture (Ravindran & Myers, 2012). While Western culture allows disabilities such as autism and impaired intellectual functioning to be viewed openly as a health challenge with possible solutions, most African cultures and similar environments are not yet completely open to discussing these disabilities (Kromberg et al., 2008; Ravindran & Myers, 2012). In Nigeria, as in most African countries, if parents observe developmental delays, the children are hidden, alternative treatments and local solutions are used, or help is sought only during adolescence. As such, decisions on whether to seek help, where to seek help, how to seek help, resources available and interventions are affected by culture (Aguwa, Carrasco, Odongo & Riblet, 2022; WHO, 2023; Kamau, 2017; Bedford & Sharkey, 2014; Ravindran & Myers, 2012). Since Western culture allows for more proactive measures in intellectual disability and autism, there is continuous development and improvement of screening, allowing for the provision of supportive services (Towle & Patrick, 2016; McKenzie & Paxton, 2012, McKenzie et al., 2012a; Nijman et al., 2018; Ravindran & Myers, 2012). Such services include screening for these disorders, diagnosing and providing intervention.

Persons with either intellectual disability or autism may place enormous financial burdens and strain on their families, especially where there is a lack of government intervention and adequate healthcare facilities, as in Nigeria (Baird et al., 2006; Gureje & Lasebikan, 2006). Presently in Nigeria, the practice is for medical professionals to "diagnose" both intellectual disability and autism based on the DSM-V or ICD-11 criteria without the use of an acceptable gold standard tool. However, diagnosis using a gold-standard tool can be time-consuming, costly, and requires training, which is generally limited in Africa and Nigeria, the focus of this thesis. For example, the Autism Diagnostic Observation Schedule (ADOS-2), Autism Diagnostic Interview-Revised (ADI-R), or the Weschler intelligence scales require the purchase of expensive equipment and explicit training to administer the tests. Screening, on the other hand, can be quick and effective, with minimal training depending on the tool (Iragorri & Spackman, 2018), thus making the availability of short screeners a necessity in Africa and Nigeria. Also, some professionals who currently screen typically use

readily accessible measures, such as the M-CHAT, designed for young children and not yet validated for the Nigerian population. Considering that most children are hidden away till later in life, it seems that these kinds of early screening tests are not useful, as many children who are not diagnosed early in life within Nigeria are brought to the attention of professionals around the onset of adolescence (Franz et al., 2017). Adolescence is when teenagers spend more time away from the family home. Adolescents and young people are those aged 11 to 26 years, this period being consistent with the critical period of brain maturation associated with development during adolescence (Sawyer et al., 2012; Sawyer, Azzopardi, Wickremarathne & Patton 2018).

There is a marked absence of well-developed screening tools for use with adolescents among professionals and services in Nigeria. Overall, screening for intellectual disability or autism in individuals over 36 months of age in Nigeria requires that a validated and reliable measure becomes accessible to front-line professionals such as nurses, carers, family doctors, and primary health care services. Doing so can reduce costs and time as only those with an increased probability of having the condition will progress to a full diagnostic assessment. While it is necessary to save time and cost, caution and attention are required as the reliability and validity of the tool cannot be traded (Watson et al., 2007). To begin addressing this barrier and attempt to move in the same direction as the Western cultures, screening tools validated with the Nigerian population - for intellectual disability and autism would be valuable.

2.5 Autism Screening Tools

In Chapter 1, various theories of autism were highlighted, with none singled out as addressing all the domains of autism. The availability of numerous screening tools may reflect age categorisations and should reflect all theories. Previous measures were developed based on earlier versions of diagnostic criteria. With the revised DSM-5 and ICD-11, some of the older tools have been revised. Due to the complexity and heterogeneity of autism, a range of tools is required to obtain relevant information.

Autism screening tools include checklists, questionnaires, parent interview forms, self-rating scales, and observational reports. The questionnaires, checklists and parent forms are usually questions with dichotomous ‘yes/no’ responses. Responses are tallied, and scores are compared to a suggested cut-off for autism. Some of the commonly used ones are the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, Barton & Green,

1999), Social Communication Questionnaire (SCQ; Berument, Rutter, Lord, Pickles & Bailey, 1999), Screening Tool for Autism in Toddlers (STAT; Stone, Coonrod & Ousley, 2000), and Social Responsiveness Scale (SRS; Constantino, Davis, Todd, Schindler, Gross, et al., 2003). Of these, only the SCQ covers a wide age range; 4 years to adulthood. Some professionals consider that a diagnosis of autism can be made when the requisite DSM-5 symptoms are present and other disorders are adequately ruled out. However, most experts consider that the gold standard tools for the diagnosis of autism are the parent interviews – ADI-R (Becker et al., 2012; Kleinman et al., 2008) or the Diagnostic Interview for Social and Communication disorders (DISCO; Billstedt, 2007; Charman & Gotham, 2013; Leekam, Nieto, Libby, Wing & Gould, 2007) and the ADOS (Kleinman et al., 2008; Gotham, Risi, Pickles, & Lord, 2007; Luyster et al., 2009).

2.6 Intellectual Disability Screening Tools

Because a multicomponent assessment determines intellectual disability, different tools are used for screening. Most of the tools are administered directly to the individual. A few level 2 tools, such as the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q; McKenzie et al., 2012), Screener for Intelligence and Learning Disabilities (SCIL; Nijman et al., 2018), Hayes Ability Screening Index (HASI; Hayes, 2002), are available for use with adolescents. However, most of these tools have not been validated outside of the original Western environments. Other standardised measures developed to determine levels of intellectual functioning measured by IQ scores and adaptive behaviour are typically administered for diagnostic purposes. Such measures include the gold standard tools such as the Weschler scales and Vineland Adaptive Behavior Scales. These are lengthy and require specific training. The “Ten Questions” (TQ), on the other hand, is a disability screening tool which has been used widely in developing countries (Stein, Durkin & Belmont, 1986; Kakooza-Mwesige et al., 2014; Durkin, Hasan & Hasan, 1995). The TQ was primarily designed as a stop-gap screening tool for numerous kinds of impairment in children aged 2 – 9 years old, including intellectual disabilities, and has been used to estimate prevalence within low-income and low-resource countries. The TQ concerns cognitive skills, motor skills, hearing, epilepsy, and vision problems.

Stein et al. (1986) used the TQ measure as a screening tool in the first stage of their prevalence study across several countries to identify children with moderate to profound intellectual disabilities. Study samples were from eight countries (India, Philippines, Bangladesh, Sri Lanka, Malaysia, Pakistan, Brazil, and Zambia). Intellectual disabilities

were classified as an intelligence quotient less than or equal to 55 (IQ ≤ 55). Stein et al. (1986) did not use standardised IQ tests, such as the Weschler scales, for classifying children with intellectual disabilities. Assessment of the 'disability' as defined by Stein and colleagues was somewhat arbitrary and based on clinical judgment as a "satisfactory standard method to measure this component across cultures (for instance, with IQ tests) has not been devised" (p. 11).

Furthermore, Stein et al. (1986) acknowledged that there could have been differences between professionals in their definitions of intellectual disability. No specific figures were reported for sensitivity and specificity. However, the team reported that most participants with intellectual disabilities were probably identified as well as other children with other conditions and IQs greater than 55. These results from Stein et al. (1986) do not seem adequate to judge the psychometric properties of the TQ. Also, any consideration of cultural issues was not documented. Two other studies, Mung'ala-Odera et al. (2004) in Kenya and Kakooza-Mwesige et al. (2014) in Uganda, used the TQ with children in their early years. Kakooza-Mwesige and colleagues screened 1,169 Ugandan children between the ages of 2 and 9 years using an adapted version of the TQ, which included 13 additional questions about autism. Questions about autism covered the three criteria: qualitative impairment in social interaction, qualitative impairments in communication and restricted repetitive and stereotypical behaviours. This adapted version of the TQ was called the 23Q. The authors reported a high negative predictive value (.90) and specificity (.90), with a very low positive predictive value (.22) and sensitivity (.52) for participants with autism, implying a high rate of false negatives. As such Kakooza-Mwesige et al. (2014) concluded that neither the TQ nor the 23Q met the criteria as useful screening tools for autism. While the TQ has helped identify children with specific and severe disabilities, its appropriateness for detecting more complex and hidden disabilities, such as mild intellectual disabilities or autism, is unclear (Olusanya & Okolo, 2006; Durkin, 2001). Suggestions have been made that continued use of the TQ in Africa or LMICs may undermine effective screening and early intervention efforts because of continued reliance on parent's and caregiver's reports without organised efforts targeted at early detection of impairments for timely, efficient, and effective intervention (Olusanya & Okolo, 2006).

2.7 Psychometrics and Measurement Properties of Screening Tools

Psychometrics refers to the theory and technique of measuring mental processes, aimed at explaining the meaning of responses provided by participants in a study (Pasquali, 2009;

Wallsten & Budescu, 1983). Psychometric and measurement properties also refer to diagnostic or screening tools' reliability, internal consistency and validity (Wallsten & Budescu, 1983). Steven (1946) defined measurement as "assigning numbers to objects and events in accordance with given rules" (p. 677). In psychology, assigning numbers to intangible constructs is difficult due to subjectivity and challenges defining the attributes of interest as they are abstract and sometimes unobservable (Wallsten & Budescu, 1983). Therefore, certain observable aspects of an individual's response may indicate such concepts' existence (Borsboom, 2005; DeVellis, 2006). For example, intelligence or intellectual functioning as a construct is abstract and unobservable, but tests have been developed for assessing observable behaviours which can be measured that reflect intelligent functioning. The ability to assign numbers has invariably led to different scales (nominal, ordinal, interval, or ratio), measurements, and, in turn, different rules or theories (Steven, 1946; Wallsten & Budescu, 1983). Therefore, the rules for number assignment, the structure of screening tools and the statistical operations applicable to measurements must be explicitly defined (Wallsten & Budescu, 1983). Different measurement models were developed to address this – classical test theory (CTT), item response theory (IRT), generalisability theory (G theory), representational measurement theory (RMT), and latent variable (LV). IRT and G theory will be briefly described, and CTT, the method used for this thesis, will be discussed further.

2.7.1 Classical Test Theory

CTT, also known as the 'True Score Model', describes psychometric procedures used to examine the sum of item responses in a screening tool. The score obtained by an individual on any given test or the score from a screening tool is given meaning by comparing the score to a predetermined cut-off (DeVellis, 2006). For instance, an autism screening tool with a predetermined cut-off score of 15 means that any individual who scores 15 or over on that tool will meet the criteria for autism. CTT also posits that an individual's test score comprises the true score and measurement error. However, the measurement error can be resolved through test-rest reliability or parallel forms (Borsboom, 2005; Mokinkk et al., 2018a; Hambleton & Jones, 1993). CTT has been criticised for various reasons, including the assumption that individual and assessment characteristics cannot be separated and the assumption that the standard error of measurement is the same for all individuals. Other criticisms are that CTT is tool-oriented rather than item-oriented, and the model's assumptions are easily met (Hambledon & Jones, 1993; Hambleton, Swaminathan &

Rogers, 1991). A further argument by Hambledon et al. (1991) is that "scores on any test are unequally precise measures for examinees of different ability, thus making the assumption of equal errors of measurement for all examinees implausible" (p. 4). CTT recognises that there is no psychological measurement void of error and that such error is random (DeVellis, 2006). However, "random errors will tend to average out as the number of observations increases" (Borsboom, 2005, p. 15). Thus, because the error is random, "their mean is zero … when all sources of error are combined, they should cancel each other out and have little or no effect on the item mean" (DeVellis, 2006, p. S51; Raykov & Marcoulides, 2011). The concept of measurement error from the CTT perspective makes sense considering the characteristics of the population studied in this thesis – persons with autism and intellectual disability. The difference between their screening scores and ability can be considered an error, given that autism is heterogeneous. Although CTT has been criticised for focusing on tool quality and its validity rather than on the items, this was considered a strength (Pasquali, 2009) in this thesis. The focus of this thesis was not on the development of new screening tools but on the validation of existing tools. Also, most existing screening tools are based on the CTT model, with most statistical packages built to compute CTT outcomes. Operations such as factor analysis and coefficient alpha, the ease of use amongst researchers, and CTT's popularity in the field of psychology despite the emergence of newer models (DeVellis, 2006; Mokinkk et al., 2018a; Rusch, Lowry, Mair & Treiblmaier, 2017) contributed to the use in this thesis.

Tools vary in their psychometric properties and feasibility. Screening tools can be used across all developmental domains or limited to a specific condition such as intellectual disability or autism (Gladstone, 2005). Whatever the focus of the screening tool, the psychometric principles underlying their construction are the same. Psychometric properties include sensitivity, specificity, convergent validity, discriminant validity, construct validity, reliability, internal consistency, cross-cultural validity, and structural validity. Depending on the exact purpose of an assessment, specific properties of the screening tool can be assessed. Following are the descriptions of a few of these properties relevant to this research.

- Sensitivity: the percentage of people with actual disabilities correctly identified by a screening test as having a disability (Glascoe, 2005).

- Specificity: the percentage of people without disabilities correctly identified by negative findings on screening tests (Glascoe, 2005).
- Positive Predictive Value: the probability that the condition is present given a positive test result (Wong & Lim, 2011).
- Negative Predictive Value: the probability that the condition is absent given a negative test result (Wong & Lim, 2011).
- Internal Consistency: the extent of the interrelatedness among the items (how well the items measure the same construct) (Mokkink et al., 2018a).
- Cross-cultural Validity*:* the extent to which the performance of the items on a translated or culturally adapted tool adequately reflects the performance of the items of the original version (Mokkink et al., 2018a).
- Structural Validity: the extent to which the scores of a measure are an adequate reflection of the dimensionality of the construct to be measured, typically assessed by factor analysis (Mokkink et al., 2018a).
- Construct Validity: the perceived overall validity of the measurement, typically measured by correlation coefficient or factor analysis (Asunta, Viholainen, Ahonen & Rintala, 2019).
- Convergent Validity: the extent to which two measures of a construct that theoretically should be related are related (Asunta et al., 2019).
- Face validity is the extent to which one or more individuals think a measure appears to cover the concept it claims to measure (Asunta et al., 2019).
- Discriminative Validity: verifies that measures that should not be related are not related, typically measured by correlation coefficient (Asunta et al., 2019).

Glascoe (2005) states that "the accuracy of a screening test is defined by its sensitivity, specificity, and positive predictive value" (p. 174); however, while this may be essentially correct, the other measures listed above are also valuable.

Under DSM-IV-TR and ICD-10, the diagnostic criteria for autism required impairment in each of the three domains: social interactions, social communication, and restricted and repetitive interests and behaviour, with onset by 3 years of age (Volkmar et al., 2014). Therefore, most tools and practice guidelines were developed to identify these symptoms in early childhood (Ozonoff, Goodlin-Jones & Solomon, 2005). Since developmental profiles differ in and for individuals as they grow older, observed symptoms may present differently; therefore, if a tool designed for very young children is used with adolescents, the psychometrics must be re-evaluated. For instance, an individual with deficits in the social communication and interaction domain as a young child may acquire some social skills as they grow older but may still not be typical of others of his/her age. Also, the psychometrics must be re-assessed when the screening tool will be used in a population different than that used during the initial development.

Therefore, before stating that a tool has excellent psychometric properties, meaning a measure is reliable and valid, it must be evaluated extensively. The accuracy of screening tools is vital, and Glascoe (2005) recommends that the sensitivity, or the true positive rate, should be between 70-80%, while specificity, or the true negative rate, should be at least 80%. In order to confirm the screening tool's validity, the outcomes are compared to those of a generally acceptable gold-standard diagnostic tool (Maxim, Niebo & Utell, 2014).

2.7.1.1 Cross-Cultural Appropriateness

Screening tools require validation when used outside the environment and population with whom they were developed, referred to as cross-cultural validation. Sometimes the process involves translation or simply utilising the tools in a different setting or with different genders or samples (Mokinkk et al., 2018a). Mokinkk et al. (2018a) define cross-cultural validity as "the degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the instrument" (p. 51). While efforts are made to ensure that measurement outcomes are replicated in the new environment, Küçükdeveci, Sahin, Ataman, Griffiths & Tennant (2004) have suggested a broader approach. Küçükdeveci et al. (2004) recommend that when the intent is to use adapted versions of the tool in a different context, then the "probability of a patient in one country affirming an item (in the dichotomous case) will be the same as the probability of a patient in another country affirming the item, given that they are both at the same level of the trait or construct being measured" (p.14). This latter submission was more relevant to the SCIL, translated from Dutch to English for this thesis.

In addition, cultural validity includes using contextualised examples and language within the sampled population by ensuring that the "instruments are grounded" within the specific culture to be considered valid (Jadhav, 2009).

Most screening tools in existence have been validated in the West, but evidence for their validation in Africa is scant (Soto et al., 2015; Van der Linde, Swanepoel, Glascoe, Louw & Vinck, 2015). Another consideration is the adaptation of measures for use outside of the original design environment. A robust screening tool should be culturally sensitive and useable with multiple populations (Van der Linde et al., 2015). Given that almost all the measures were developed within Western countries, issues regarding cultural sensitivity and the feasibility of using these screening tools in their original format with the African populace need investigating. Screening tools developed in high-income environments do not necessarily consider the application and understanding of the terminology in other environments. Screening results and reliability can be affected where the language of the screening tool differs in application or understanding (Soto et al., 2015).

In Africa, some studies that measured developmental milestones and disabilities utilised screening tools developed in the West (Oshodi et al., 2016; Koura et al., 2013; Jinabhai et al., 2004). For example, Oshodi et al. (2016) used the Modified Checklist for Autism in Toddlers (M-CHAT) and the Childhood Autism Spectrum Test (CAST) in a Nigerian urban setting where language and terminology were not barriers, thereby eliminating the need for translation and found that 34.5% of the participants were autistic. However, they reported no psychometric properties of the tools. Jinabhai et al. (2004) adapted and substituted examples in both the Auditory Verbal Learning Test (AVLT) and Young's Group Mathematics Test (GMT) with more familiar items for Zulu participants. The AVLT instructions were given in Zulu with the items' turkey' and 'ranger' replaced with the more culturally familiar Zulu words 'chicken' and 'herdboy', respectively. In addition, Jinabhai et al. (2004) made considerably more adaptations to the GMT and administered the test in Zulu. The adaptations centred on the change of words and examples to more familiar items such as 'tarts' to 'cakes', 'marbles' to 'balls', 'engine' to 'truck', and the names 'Dick and Jim' were changed to 'Sipho and Thembi'. Jinabhai et al. (2004) found that the test scores and the mean scores of the Zulu children in their study were lower than those reported in other studies. Therefore, they suggested the development of more appropriate test instruments for South African conditions. Meanwhile, Koura et al. (2013) adopted a rigorous translation model to translate and adapt the Mullen Scales of Early Learning

(MSEL) used in their study from English to French while the parents' instruction was translated into Fon, the local language. They reported a correlation score for inter-rater reliability between the gold standard interviewer (the assessor who administered the most tests) and other assessors in the range of 94% to 100%, with a mean correlation score of 98%.

Some other studies highlight the importance of language and terminology in translated versions. Wild et al. (2012) translated the CBCL into six languages Korean, Hebrew, Spanish, Kannada, and Malayalam. In the Malayalam version, several cultural adaptations, such as changing the "milk delivery" to a more familiar job and giving different examples of sports and hobbies, were needed, while in the Hebrew version, two sexually related items were removed from the measure. Koura et al. (2013) also used the "Ten Questions" (TQ) to screen for disabilities and collect cognitive development information for their participants; however, the items were not translated into other languages.

2.7.2 Item Response Theory

Item response theory is one of the two extensions of CTT for examining measurement properties in assessment tools. The IRT, a complementary model to the CTT rather than a superior model (Jabrayilov, Emons & Sijtsma, 2016), was developed to address some of the gaps identified by psychometricians in the CTT (Hambledon et al., 1991). One such is focusing on item quality in assessment tools rather than the quality of the tool itself. That is, IRT focuses on item validity. It also measures the ability of individuals on each item in the assessment tool/test by assuming that each item (test question) has a unique characteristic, such as difficulty and discriminating power. However, Jabrayilov et al. (2016) reported that initial results revealed that IRT performed better than CTT in individual change detection for tools with at least 20 items, while CTT was better suited for tools with fewer questions.

This thesis employed two screening tools, the SCQ and the SCIL, as will be discussed in Chapter 4. The SCIL is a relatively new and short tool with 14 questions, and the CTT model was used for psychometric analysis, given the advantage over IRT for short tools. The SCQ is an older tool with 40 questions and has been used widely in research. This research did not include psychometric analysis using IRT as there is sufficient positive evidence for the psychometric properties of the SCQ Lifetime form used in this thesis based on the IRT model (Wei, Chesnut, Barnard-Brak & Richman, 2015; Karaminis & Stavrakaki,

2022). Additionally, the participants numbers were insufficient to attempt an IRT on the SCIL but met the size criteria for CTT (Brown & Abdulnabi, 2017).

2.7.3 Generalisability Theory

G theory, an extension of the CTT, is used to evaluate the reliability of measurements and pinpoint the sources of measurement error (Webb & Shavelson, 2005). In G Theory, sources of variation are termed 'facets', analogous to factors in variance analysis. Facets may include persons, items/questions or other variables and can be random or fixed. In psychological assessments, individuals are the object of interest. G Theory focuses on the components of variance associated with the object of measurement, with the facets and their interactions (Di Nocera, Ferlazzo & Borghi, 2001). The component of variance associated with a facet reflects how much that facet contributes to the measurement error. In G Theory, variation among individuals represents actual differences, and variations associated with the facets and their interactions represent the errors that affect the measure (Brennan, 2010; Shavelson, Webb & Rowley, 1989). For example, the autism screening tool can be administered to a participant by persons A and B, then scored by person C. The G Theory assumes that measurement errors arise from inconsistencies in administration, scoring or other variables and incorporates each potential error source in the measurement model. Additionally, the G Theory allows researchers to decide on the facets of interest while all other facets are considered sources of measurement error (Brennan, 2010; Shavelson, Webb & Rowley, 1989).

While there are similarities between the CTT and G Theories, there are differences. Both theories define true (or universe) scores as an expected value of observed scores, incorporate random measurement errors, and have well-defined notions of reliability. Their differences lie in the definition and source of errors; CTT has a single source, and G Theory recognises multiple sources making it more complex to analyse. The complexity of the G Theory has deterred researchers who prefer the simplicity of the CTT model in analyses of screening outcomes (Brennan, 2010). Another difference between the CTT and G Theory is in making criterion-referenced versus norm-referenced decisions. For example, classifying a participant as autistic or determining the general level of intellectual functioning based on the cut-off score (criterion-referenced) versus comparing the individual's score to that of peers (norm-referenced) or comparing current versus previous scores on the same test (Hintze, Owen, Shapiro & Daly III, 2000; Kunnan, 1992). For criterion-referenced

assessments, CTT is the preferred analysis method (Sawaki, 2016). Since the focus of this thesis was primarily criterion-referenced, G Theory was not explored.

2.8 Summary

Standardisation of measures establishes uniform procedures for assessment administration and scoring so that any conclusions derived from the evaluation are as objective as possible, valid, and reliable. Different models for examining the measurement properties of the screening tools have been examined in this chapter, and the most suitable option has been identified. No study has evaluated an existing screening tool for use in Nigeria, for autism or intellectual disabilities, especially among adolescents. Therefore, given the absence of validated screening tools, broad-age lifetime screening tools, which cover most, if not all, of the autism and intellectual disability symptoms, will be identified and validated for use with Nigerian adolescents. A systematic review was completed to identify such tools, which will be discussed in Chapter 4.

Chapter 3. The Nigerian Context and the Plan for the Screening Studies

In Chapter 1, an overview of autism and intellectual disability was presented. In Chapter 2, the state of screening for intellectual disability and autism within Africa, along with the basis of screening and the psychometric requirements, were highlighted. In this chapter, the gaps in the research literature and the challenges with screening for autism and intellectual disability in Nigeria will be highlighted. The different studies aimed at contributing to resolving the identified gaps, specifically in Nigeria, will be outlined.

3.1 Introduction

There is a paucity of epidemiological studies of intellectual disability and autism, with minimal information on screening, diagnosis, and interventions, in LMIC countries. Nigeria in West Africa falls in this category. Given that screening and early diagnosis positively impact intervention outcomes (Koegel, Koegel, Ashbaugh & Bradshaw, 2014), having the diagnostic and screening tools is vital. From previous research (Nwokolo, 2017), the barriers to assessment and intervention for people with autism in Nigeria were identified as a lack of reliable tools for screening and diagnosing, inadequately trained personnel to administer the tools and/or trained personnel without access to tools validated for the Nigerian population. Also, the culture of silence about disabilities, fear of stigmatisation, shame, diverse religious beliefs, and individuals not seeing the value in screening or getting a proper diagnosis for autism, intellectual disability, or other developmental disorders was problematic. Sango (2017), in her review, presented a similar overview of the situation of people living with intellectual disability in Nigeria, noting the lack of individually administered measures, inadequate and inaccurate diagnosis, inconsistent assessment processes, lack of adequate screening tools, lack of expertise to provide services to this population, multi-ethnicity and lack of government support as challenges.

There are over 200 ethnic groups and 500 indigenous languages in Nigeria (CIA, 2023; Kayser-Jones, Abu-Saad & Akinnaso, 1982), with Igbo, Hausa, and Yoruba being the dominant languages and groups. Health administration in the country is on three levels – the federal government, state government, and local government. The Federal Ministry of Health (FMH) is responsible for formulating health policies and safeguarding quality control. The government is also in charge of the public health infrastructure. In the health

sector, there is an overwhelming demand for professionals and infrastructures, especially in mental health and developmental disabilities. Amakom (2012) reported that more than 58% – 70% of poor families made use of public healthcare facilities, while between 14% – 29% of affluent families use the same facilities. Similarly, results from that study showed that over 70% – 80% of the richest families patronise private healthcare facilities when compared to between 13% – 31% of extremely poor families. While there are psychiatrists, psychologists, and paediatricians in the public health sector, they are not adequately equipped to screen for autism or intellectual disabilities. Qualified professionals within the private sector with resources to screen for autism or intellectual disability are limited and overburdened. In addition to Amakom's (2012) findings, there is the burden of "bad" governance in the health sector (Adeloye et al., 2017) and underutilisation of the facilities by the poor due to lack of finance (O'Donnell, 2007), religious and individual beliefs.

3.1.1 Context of the Research

Although there are some trained psychologists, psychiatrists, behaviour analysts, physicians and other healthcare professionals in Nigeria, limited efforts have been made towards screening and diagnosing children with intellectual disabilities or autism. The few clinicians and healthcare practitioners who screen patients for these conditions mostly use tools which have not been validated with the Nigerian population. Occasionally, non-age-appropriate screening tools are used as they are readily available online. Thus, a sizable number of individuals, especially adolescents, are either misdiagnosed, remain in their parent's homes or in government-owned psychiatric hospitals or remand homes where they receive limited to substandard interventions for their conditions or are abandoned by parents and society in general (Atilola, Omigbodun, Bella-Awusah, Lagunju & Igbeneghu, 2014). Furthermore, peer-reviewed research work in these areas remains extremely minimal.

Diagnosing intellectual disability or autism requires substantial training, competency, and experience with persons with an intellectual disability or autism (Rogers, Goddard, Hill, Henry & Crane, 2016; Volker & Lopata, 2008). Diagnosing intellectual disability or autism also requires the collaborative effort of a multidisciplinary team (Rogers et al., 2016; Yates & Le Couteur, 2016). Limited professionals have sufficient training and capacity to administer gold-standard diagnostic tools in Nigeria. Where they exist, they are predominantly located in urban and large commercial cities such as Lagos and Abuja. As such, persons who are suspected to have either an intellectual disability or autism must wait long periods before identifying or accessing a professional or go undiagnosed. Individuals'

resident outside of the large cities and urban areas are in worse situations (Bakare & Munir, 2011; Bakare et al., 2008). In a country of over 200 million people (World Bank, n.d.), with approximately 109 million persons under the age of 18 (UNICEF, 2014), the professional-to-children and adolescent ratio are overwhelmingly low. Whereas diagnosing intellectual disability or autism requires extensive training, competency, and collaborative efforts, screening for either condition does not. Individuals such as teachers, frontline health practitioners, public and primary health care workers with minimal training can administer screening tools. Therefore, having screening instruments that do not require advanced-level training can enable more professionals and paraprofessionals to detect individuals who may have autism or intellectual disability in their environment while allowing the available professionals to focus efforts on diagnosing and intervention. Additionally, data gathering on the prevalence and incidence rates of these disabilities can begin.

3.1.2 The Gap

There are very few published studies investigating the utility and validation of screening and diagnostic instruments for both intellectual disabilities and autism within the African continent (Franz et al., 2017). Where studies have been done, the focus has been on younger children (Marlow, Servili & Tomlinson, 2019; Bozalek, 2013). A recent study by Awadu (2021) assessed the psychometric properties of the SCQ and Social Responsiveness Scale-Second Edition (SRS-2) against the TQ among 4 to 18-year-olds in Uganda and concluded that the SCQ and SRS could be used with the Ugandan population. Attempts to assess the level of general intellectual functioning amongst adolescents using standardised tests (e.g., Slosson's Intelligence Test) have been made, but questions remain about the robustness of attempts to validate these tools (Atilola et al., 2014). Out of the 67 adolescent participants suspected to have an intellectual disability in the Atilola et al. (2014) study, 46.7% were flagged as having intellectual disabilities of varying degrees, including mild intellectual disability. Other neurological disorders such as epilepsy, slurred speech, and dyskinesia in children as co-morbidities with intellectual disability were examined within the Atilola et al. study. Concerning autism, Oshodi et al. (2016), in their community-based study, utilised the M-CHAT and the CAST as screening tools. However, there was no evidence of validating either tool for the Nigerian population. Given the insufficient evidence for validating any of the tools mentioned above, little is known about the tools' performance. A common factor in these studies is the lack of access to accepted gold-standard tools for assessing convergent validity.

The current research aims to address the gap in validating intellectual disability and autism screening tools for use in the Nigerian context. Such tools must be robust enough for use amongst Nigerian adolescents. To do so, a series of studies leading to the validation of specific screening tools for autism and intellectual disability, which are time efficient and require little to no training to administer, was undertaken.

3.1.3 Research Questions

The main research questions are:

1. Is there a screening tool for intellectual disability which requires little or no training that can be used amongst Nigerian adolescents?
2. Is there a screening tool for autism which requires little or no training that can be used amongst Nigerian adolescents?

These questions will be answered through the different studies outlined below.

3.2 Thesis Outline

This research consists of three studies. The first study consisted of a systematic review to identify short screening tools for detecting intellectual disabilities and autism in older children and young people aged 11 to 26 years. Then the published psychometric properties of these tools and the appropriateness of using these tools across a range of cultures were evaluated.

In study 2, using a group of experts (consensus group), the face, content, and cultural validity of the screening tools identified through the systematic review were examined, and adaptations were made as required. The selected screening tools from the focus group study were the SCQ for autism and the SCIL for intellectual disability.

The third and final study was in two parts, validation for autism and intellectual disability screening tools. The autism screening tool and the intellectual disability screening tool were administered to participants, together with a full diagnostic assessment. Data gathered with the screening tools were analysed to evaluate aspects of its internal consistency, convergent validity, discriminant validity and other relevant properties. The third study will be discussed in two parts – the validation of the intellectual disability tool (3a below) and the validation of the autism screening tool (3b below). These studies will be described in further detail in the sections to follow.

3.2.1 Study 1 – Systematic Review

A systematic review was first conducted to identify and evaluate appropriate and available screening tools for intellectual disability and autism in adolescents and to determine any gaps in research on and availability of screening tools. The systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Page et al, 2021) guidance. To provide transparency of the review process and avoid duplication of the study, the review was registered with Research Registry (https://www.researchregistry.com/ - Registration Code: reviewregistry798). Research Registry is an international database for registering all types of research studies, such as case reports, observational and interventional studies, systematic reviews and meta-analyses. A quality appraisal of the included studies was conducted using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) risk of bias checklist (Terwee et al., 2018a; Prinsen et al., 2018; Mokkink et al., 2018b) and the manual as guides. The COSMIN manual was developed for the systematic review of patient-reported outcome measures (PROMs). Although this review did not consider PROMs, the COSMIN checklist was adopted due to its robustness. The relevant aspects of the COSMIN for this review include its usefulness for assessing the methodological quality of studies, development and design of measurement tools, psychometric properties, and cultural validity. The application of the COSMIN to the systematic review is discussed further in Chapter 4.

The aims of the systematic review were to:

1) describe and critically appraise short screening tools for the detection of intellectual disabilities and autism in children and young people aged 11 to 26 years,
2) consider the psychometric properties of these tools, and
3) consider the appropriateness of using these tools across a range of cultures.

This systematic review has been published (Nwokolo et al., 2022; Appendix 20).

3.2.2 Study 2 – Focus Group

During this stage, the screening tools identified through the systematic review were presented to a group of experts using the nominal group technique (NGT), also known as a focus group. The focus group technique was chosen to due to its extensive use in studies for similar decision-making (Humphrey-Murto, Varpio, Gonsalves & Wood, 2017;

International Test Commission, 2017). The process is based on the notion that valid, accurate and reliable evaluation is best achieved by consulting a team of experts and stakeholders. Achieving accurate and reliable assessment is assumed to be achievable through the group (Humphrey-Murto et al., 2017), and consensus methods have been used in education for curriculum development (O'Neil & Jackson, 1983), as well as in medical and health research (Humphrey-Murto et al., 2017; Van de Ven & Delbecq, 1972). Several studies support the use of the consensus group methods in developing items for measurement tools, developing clinical guidelines, and deciding on components of new or revised curricula (Van de Ven & Delbecq, 1972; Murphy et al., 1998; Humphrey-Murto et al., 2017).

The aims of this study were to:

1) consider the face, content, and cultural validity of our chosen screening tools and
2) make recommended adaptations for use with Nigerian adolescents using a consensus group methodology. Details of the procedure are in Chapter 5.

This focus group study has been submitted for publication (Nwokolo, Murphy, Mensink, Moonen & Langdon, submitted; Appendix 20).

3.2.3 Study 3a – Preliminary Testing of the English Version of the Screener for Intelligence and Learning Disabilities

Following the agreement and recommendation of the focus group, the Screener for Intelligence and Learning Disabilities (SCIL) was tested with the Nigerian adolescent population to screen for intellectual disability. For the validation study, the following questions were asked.

1. In the Nigerian context, how will the psychometric properties of the English SCIL compare with the original Dutch version?
2. Within the Nigerian context, what will be the appropriate cut-off score for identifying persons suspected of intellectual disability using the English SCIL?
3. Are all of the original factors of the SCIL relevant in the Nigerian context, or will there be a need for dimension reduction?
4. Will the SCIL be a valid screening tool for use in the Nigerian context?

Thus, the aims of this study were to

a) examine the component structure of the SCIL and reduce dimensions as required,

b) examine the internal consistency, discriminant, and convergent validity of the SCIL,

c) derive an appropriate cut-off score based upon sensitivity and specificity and

d) derive the positive and negative predictive values.

Details of the validation are in Chapter 6. The study has been submitted for publication (Nwokolo et al., submitted; Appendix 21).

3.2.4 Study 3b – Validating the Social Communication Questionnaire

The SCQ is a well-developed and widely used screening tool for autism. It was recommended in study 2 by the consensus group, as an appropriate tool for screening for autism, with minor changes of wording in some items (for cultural reasons). To validate the SCQ, the following research questions were asked:

1. In the Nigerian context, will the psychometric properties of the SCQ meet those reported in the manual or previous validation studies?
2. Within the Nigerian context, what will be the appropriate cut-off score for identifying persons suspected of autism using the SCQ?
3. Are all original factors of the SCQ relevant in the Nigerian context?
4. Will the SCQ be a valid screening tool for use in the Nigerian context?

To ascertain these, the aims of study 3b were to

a) validate the structure of the SCQ in the Nigerian population using confirmatory factor analysis (CFA)

b) examine the internal consistency, discriminant, and convergent validity of the SCQ

c) derive an appropriate cut-off score based upon sensitivity and specificity and

d) derive the positive and negative predictive values.

Details of the validation are in Chapter 7. The study has been submitted for publication (Nwokolo, Murphy & Langdon, submitted; Appendix 22).

3.3 Ethics

The nature and scope of the research necessitated a multi-tiered ethical approval. The application of each approval is described as applicable in Chapters 4 through 7.

3.3.1 Tizard Centre Ethical Approval

Ethical approval was sought from the Tizard Ethical Committee for the entire research in February 2019. Following amendments to the information sheets and consent forms, approval was given in July 2019 (Appendix 1).

3.3.2 Online Training in Research Ethics

Applications and presentations were made to various organisations, including the Centre for Autism and Neuro-developmental Disorders (CANDDO) and the Federal Neuropsychiatric Hospital, Yaba, Lagos (FNPHY), where data were to be collected. The main researcher completed the compulsory ethical training and certification required to conduct human subject studies in Nigeria (Collaborative Institutional Training Initiative – CITI program) in February 2019. (Appendix 2).

3.3.3 National Health Research Ethics Committee of Nigeria (NHREC)

The scope of the study was to cover six geopolitical zones in Nigeria for which national ethical approval was required. Based on the Tizard Ethical approval and completion of the CITI program, NHREC approval was sought and obtained in September 2019 (Appendix 3).

3.3.4 Federal Neuropsychiatric Hospital, Yaba, Lagos (FNPHY)

Following the online training, ethical approval was sought from the FNPHY and obtained in June 2020 (Appendix 4).

3.3 Summary

Within this chapter, the plan for the screening studies in the Nigerian context was outlined, as were the rationale and the key research questions. The identification and selection processes for the autism and intellectual disability screening tools were explained. The process for ethical approvals was also mentioned. The following chapters will describe the individual studies in detail.

Chapter 4. Study 1 – Systematic Review[1]

A systematic review delivers a clear and comprehensive summary of available evidence on a given subject (Mallett, Hagen-Zanker, Slater & Duvendack, 2012). Systematic reviews also help with our current understanding of a field and identifying research gaps (Gopalakrishnan & Ganeshkumar, 2013). A systematic review "should synthesise all relevant, high-quality evidence from the existing literature, reaching unbiased conclusions in a transparent and replicable manner" (White & Waddington, 2012, p. 352).

4.1 Introduction

Several studies (e.g., Eldevik, Hastings, Hughes, Jahr, Eikeseth & Cross, 2009; Steiner, Goldsmith, Snow & Chawarska, 2012; Swinkels, Dietz, van Daalen, Kerkhof, van Engeland & Buitelaar, 2006; Luckasson & Schalock, 2013; Schalock & Luckasson, 2013) have highlighted the benefits of early detection of developmental disabilities such as intellectual disabilities and autism. The benefits have included improved behavioural outcomes and family support, as well as earlier intervention. Other benefits included improved planning for educational needs and support, improved social skills, and greater cognitive and language development. These findings have emerged predominantly from Western and high-income countries with there having been very limited research from low to medium-income countries (LMICs), as indexed by the published gross national income by the United Nations (Tomlinson et al., 2014; Gladstone et al., 2010; United Nations, 2014; World Bank, 2020 & United Nations Department of Economic and Social Affairs, 2021). While the presentation of autism is the same regardless of economic status, the political climate and associated social burdens within LMICs, such as in the African countries, discourages the early detection of developmental disabilities as it is not seen as urgent, which increased the health disparities faced by this population (Emerson, 2012; Gladstone et al., 2014). The situation is similar for those with intellectual disabilities, with late identification leading to further delay of intervention.

McConachie et al. (2015) reviewed the measurement properties of some screening tools used to measure progress and outcomes in young children with autism spectrum disorder

[1] This study has been published as Nwokolo, E. U., Langdon, P. E., & Murphy, G. H. (2022). Screening for Intellectual Disabilities and/or Autism Amongst Older Children and Young Adults: a Systematic Review of Tools for Use in Africa. Review Journal of Autism and Developmental Disorders. Advance online publication. https://doi.org/10.1007/s40489-022-00342-6

aged up to 6 years. Their reviewed focused on measuring the progress and improved quality of life post intervention for participants in the West. Soto et al. (2015), in their systematic review of 21 included studies, investigated efforts towards the cultural adaptation of screening tools for use outside of the environments in which they were primarily developed. With a specific emphasis on autism spectrum disorder only, the review examined the adherence to recommended adaptation procedures and the psychometric properties of the adapted instruments. Studies about people with intellectual disabilities were excluded. The adaptation studies included in the Soto et al. review had been carried out in nineteen countries and involved ten languages. Only two of those countries are in continental Africa: Egypt, and Tunisia. The M-CHAT was used in the studies in both countries. Egypt and Tunisia are Arabic-speaking countries, and the M-CHAT was translated into Arabic. In LMICs, where resources are limited, the cost and burden of a rigorous translation and adaptation process is a barrier to acquiring reliable screening tools.

4.1.1 Studies in Africa

Recently, attempts have been made towards developing screening tools in areas such as nutrition, neurodevelopmental disabilities and mental health which are culturally sensitive for use within the African continent (Gladstone et al., 2010; Hasegawa, Ito & Yamauchi 2017; Vawda, Milburn, Steyn & Zhang, 2017). While these efforts are commendable, study populations are often limited to early childhood, with children aged 2- to 9-year-olds. The focus upon young children (2- to 5-year-olds) would allow for the implementation of interventions earlier but would miss older children (10 years and above). Relative to studies on young children, there is very little data on studies with older children and adolescents; however, studies involving adolescents are emerging (Allison, Auyeung & Baron-Cohen, 2012; Morales-Hidalgo, Hernández-Martínez, Voltas & Canals, 2017; Nijman et al., 2018). The paucity of adolescent studies is not peculiar to Africa. What appears to be unique to LMICs and Africa is the relatively low level of awareness, insufficient economic resources, insufficient numbers of professionals, and a culture of not seeking immediate help (Franz et al., 2017).

In African countries like Nigeria, Kenya, Ghana, and Uganda, awareness is growing, yet it is still common for families not to seek immediate help for individuals with autism or intellectual disability till later in life (Franz et al., 2017). It therefore remains the case that many of these children are not screened or diagnosed early in life. Such individuals are then

brought to the attention of professionals around the onset of adolescence as this is the period when teenagers begin to spend an increasing amount of time away from the family home. Adolescents and young people are those aged 11 to 26 years of age, an age range which is consistent with the critical period of brain maturation associated with development during adolescence (Sawyer et al., 2012; 2018). To identify these older children and young adults who have been missed or not diagnosed in a time-efficient and effective way, an appropriate screening tool should be available. However, there is a marked absence of well-developed screening tools for use with adolescents among professionals and services in African countries (Hirota, So, Kim, Leventhal & Epstein, 2018).

Overall, screening for either intellectual disabilities or autism in individuals in African countries requires the use of a validated and reliable measure which is accessible to front line professionals such as teachers, nurses, carers, family doctors and those who are in primary health care services. While some screening tools have been developed and validated in the West, and investigated for use in Africa, the researchers have not always compared their study results against acceptable gold standard instruments, a crucial stage in measuring the validity of tools when used in new environments. For instance, Oshodi et al. (2016) and Koura et al. (2013) obtained reasonable results from their studies. However, they did not compare their results to that of an acceptable gold standard instrument and this presents limitations. Besides selecting and validating a standardised screening instrument for use with adolescents, the tool ought to be culturally relevant for use within the African context. Through careful adaptation and translational work, screening tools developed in the West may be adopted for use in LMIC such as Nigeria, Ghana and other African countries. By doing so, some of the costs and time to develop entirely new tools can be reduced.

4.1.2 Research Aims

To identify such tools, a systematic review was completed with the following aims:

1. To describe and critically appraise short screening tools for the detection of intellectual disabilities and autism in children and young people aged 11 to 26 years.
2. To consider the psychometric properties of these tools.
3. To consider the appropriateness of using these tools across a range of cultures.

4.2 Methods

4.2.1 Search Strategy

A literature search of the following electronic databases was carried out to identify relevant studies: Academic Search Complete, MEDLINE, CINAHL Plus, PsycINFO and PsycArticles. The key search terms were 'intellectual', 'learning', and 'autism'. These key terms were then combined with disability and with screening and diagnosis. Truncated terms were used as appropriate to ensure inclusion of variations of the words. Older words used to describe people with intellectual disabilities, such as 'mentally retarded' or 'mental retardation' were also included. Titles and abstracts were the focus of the initial search. The combined search terms are found in Table 1. Backward (ancestry) searching was used to identify other papers that may be relevant from references of eligible studies. The search was done using EBSCOhost and concluded on the 22nd of June 2018. To ensure that no new studies published, or tools developed were missed, the search was updated with the same terms on the 5th of November 2020.

Table 1 – Search Terms

Population	Intell* OR Learn* OR Mental* OR Disa* OR Retard* OR Autis* OR ASD
Tool	Screen* OR Diagnos*
Combined	((TI (((intell* and disa*) OR (learn* and disa*) OR (mental* and retard*)) OR autis* OR ASD) AND (SCREEN* and DIAGNOS*) NOT (GENE* OR DEMENTIA OR SERUM OR EYE* OR ENDOCRINE OR LEUKEMIA OR SCHIZO* OR METABOLIC OR HYPOTHYRODISM OR HYPOTHYROIDISM OR CERVI* OR INSURANCE OR PSYCHIATR* OR DEPRESSION OR AUDI* OR CANCER OR LINGUISTIC OR SUBSTANCE USE OR PREVALENCE OR PARENT* OR MALNUTRITION OR TREAT* OR SELF DIAGNOS* OR PERINATAL OR PREGNAN*))) OR (AB (((intell* and disa*) OR (learn* and disa*) OR (mental* and retard*)) OR autis* OR ASD) AND (SCREEN* and DIAGNOS*) NOT (GENE* OR DEMENTIA OR SERUM OR

	EYE* OR ENDOCRINE OR HYPERKINE* OR ADHD OR LEUKEMIA OR SCHIZO* OR METABOLIC OR HYPOTHYRODISM OR HYPOTHYROIDISM OR CERVI* OR INSURANCE OR PSYCHIATR* OR DEPRESSION OR AUDI* OR CANCER OR LINGUISTIC OR SUBSTANCE USE OR PREVALENCE OR PARENT* OR MALNUTRITION OR TREAT* OR SELF DIAGNOS* OR PERINATAL OR PREGNAN*))))

4.2.2 Eligibility Criteria and Study Selection

Titles and abstracts were initially screened for inclusion based on the following criteria: (1) the article was written in English, (2) validated screening tools were used, or the study involved developing a screening tool, (3) little or no extra training was required to administer the tool, (4) the tool did not take longer than 1-hour to administer, (5) some or the majority of the participants were aged 11 years and younger than 27 years, and (6) participants in the validation sample for intellectual disabilities or autism were diagnosed by a duly qualified healthcare professional. Some articles which had multiple studies and participants across a broad age range (Baron-Cohen et al., 2006; McKenzie, Paxton, Murray, Milanesi & Murray, 2012c; Nijman et al., 2018; Deb, Dhaliwal & Roy, 2009; Kraijer & De Bildt, 2005) were included because of their relevance in at least part of their research. Studies were excluded if any of the following criteria were met: (1) the study was related to other health issues such as diabetes, cancer, visual and any other medical condition in persons with intellectual disabilities or autism, (2) the study was about linguistic and speech-related conditions, (3) the study was about developmental learning disorders/difficulties (e.g. impairments in reading or writing) (4) the tools were not for screening but diagnostic tools, (5) publications were letters, correspondences, editorials or recommendations to the editors, (6) studies had missing information on age, (7) full text was not available, and (8) additional skills or training were required to administer the tool. Due to the paucity of research with adolescents, and in order not to miss any potential screening tools, there was no restriction on publication date. Studies done in both clinical and non-clinical settings were considered. Also, the inclusion of English only articles was based on the authors language proficiency. This initial search produced over 1000 potential articles.

After removing duplicates and completing a title and abstract screen against the eligibility criteria, a total of 235 articles were retrieved for full-text screening. This led to the exclusion of a further 194 papers. Studies were excluded due to the ages of participants (n=70) or the fact that the article was not about screening tools (n=48). One of the papers excluded at this stage had no participants in the study (Al Mamun et al., 2016), another had no details of the author and the full text could not be accessed, and thirty-three were about specific learning disorders/difficulties. See Figure 1 for the PRISMA flowchart of the study selection process (Page et al., 2021). The remaining 41 studies met eligibility criteria. The eligibility criteria were applied independently by two members of the research team (EN & GM) with excellent agreement, k = 1.

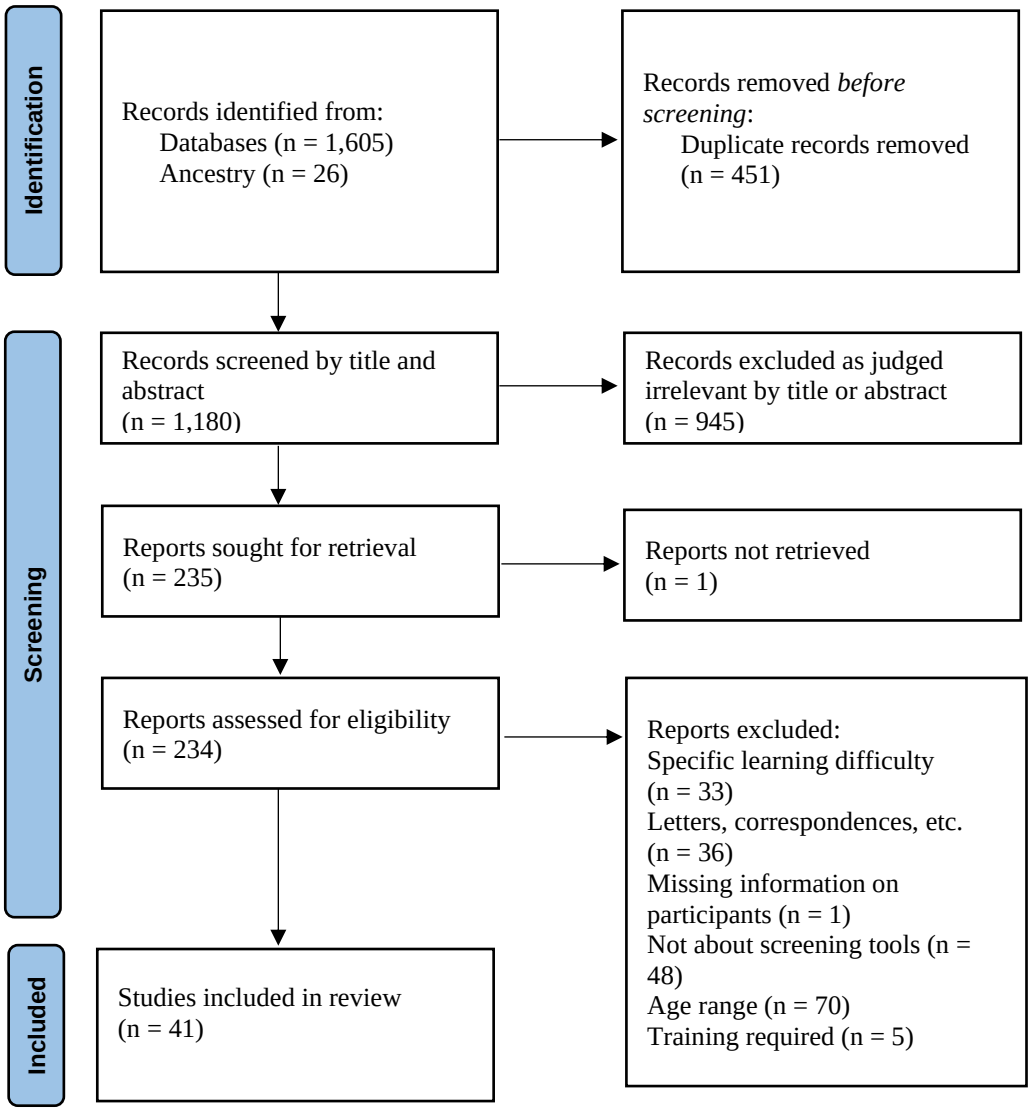


Figure 1: PRISMA flow diagram of the study selection process

4.2.3 Quality Appraisal

A quality appraisal of the included studies was conducted using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) risk of bias checklist (Terwee et al., 2018; Prinsen et al., 2018 & Mokkink et al., 2018b) and the manual as guides. The COSMIN manual was developed for the systematic review of patient-reported outcome measures (PROMs). Although this review did not consider PROMs, the COSMIN checklist was adopted due to its robustness. The relevant aspects of the COSMIN for this review include its usefulness for assessing the methodological quality of studies, development and design of measurement tools, psychometric properties, and cultural validity. The appraisal was done for all papers by EN and was independently checked by a second member of the team (PL) for 40% of the papers. Following the review of the ratings and correction of errors, the agreement was k = 1. Based on the COSMIN guidelines, the quality of included studies was rated (Tables 2 and 3). For each study, the quality was assessed based on a four-point rating system where each standard within the COSMIN box can be rated as 'very good', 'adequate', 'doubtful' or 'inadequate'. This overall rating of the study quality contributed to grading the quality of the evidence for each tool. The quality of evidence and methods were scored on a four-point rating scale, that is, sufficient, insufficient, indeterminate, or inconsistent. The overall score for quality of evidence is according to the "lowest score counts" method, and the categories used were high, moderate, low, and very low. Overall ratings for the study methodologies, quality of tool development and quality of evidence for the measurement properties using the COSMIN checklist are in Tables 4 to 7.

One key component of the COSMIN is its usefulness in evaluating cross-cultural validity of tools. Cross-cultural validity refers to "the degree to which the performance of the items on a translated or culturally adapted PROM (Patient Reported Outcome Measures) are an adequate reflection of the performance of the items of the original version of the PROM" (Prinsen et al., 2018, p. 1154). Cross-cultural validity is assessed when a tool is used with at least two different groups. Such populations could differ in language, diagnosis, gender, age groups or ethnicity (Mokkink et al., 2018a).

Additionally, the COSMIN manual suggests areas for adaptation by the review team. Some of the adaptations made for this review were related to the hypothesis testing for responsiveness (criterion validity) and construct validity. In the case of Box 9 which is

hypothesis testing for construct validity, we used it to assess the convergent validity and discriminative validity where applicable. Regarding responsiveness, Box 10a for criterion approach, we assessed the diagnostic accuracy of diagnostic tools used in the studies rather than change scores. Outcome measures of specificity and sensitivity were also assessed. For Boxes 10b and c, construct approach, studies which utilised similar measurement instruments or where the study design was between groups (children, adults, or those with and without intellectual disability or autism spectrum disorder sub-groups). Box 10d was not utilised for any studies as we did not look at interventions. Ratings of insufficient, inadequate, or doubtful were given in instances where there was insufficient information reported in the study for a higher rating as required by the COSMIN checklist. For clarification and completeness, manuals, where available, and authors of the tools were consulted for further evidence. This is discussed further in the result section.

4.2.4 Data Extraction and Synthesis

Relevant information about the aims of each included study, along with the tool used, the design, participants, time to administer the tool, and outcomes were extracted and are reported in Tables 8 and 9. The tables were arranged alphabetically by the first author, and chronologically when the first author co-authored more than one study. All included studies were quantitative.

Table 2 - Quality of studies on measurement properties for autism

STUDY & TOOL	Content validity[1]					Structural validity[1]	Internal consistency[1]	Cross-cultural validity[1]	Reliability[1]	Criterion validity[1]	Construct validity[1]		Responsiveness[1]		
	Asking patients			Asking experts											
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness						Convergent validity	Discriminative validity	Comparison with gold standard	Comparison with other instruments	Comparison between subgroups
Vrancic et al. (2002) - ADI-TSS	I	I	I	I	I	I	I	I	I	I			I		A
Morales-Hidalgo et al. (2017) - EDUTEA	I	I	I	I	I	A	V	A	V	V	A	A	V	A	V
Kraijer & de Bildt (2005) - PDD-MRS	I	I	I	D	D	A	V	A	I	V	V	V	V	V	V
Cortés et al. (2018) - EVTEA-DI (PDD-MRS)	A	A	A	A	A	V	V	A	I	V	V	V	V	V	
Heinrich et al. (2018) - DiBAS-R						I	I	A	I	V		V	V		V
Allison et al. (2012) - AQ-10						V	V	V	V	V	V	V			V
Baron-Cohen et al. (2006) - AQ-10						I	V	A	V			V			V
Booth et al. (2013) - AQ-10						V		D		V		V			V
Cederberg et al. (2018) - ASSQ								I		V		V		V	V
Kopp & Gillberg (2011) - ASSQ-REV								V	A	V	V	V	I		V
Berument et al. (1999) - SCQ						A	V	A	V	V		V	V		V

Table 2 - Quality of studies on measurement properties for autism

STUDY & TOOL	Content validity[1]					Structural validity[1]	Internal consistency[1]	Cross-cultural validity[1]	Reliability[1]	Criterion validity[1]	Construct validity[1]		Responsiveness[1]		
	Asking patients			Asking experts											
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness						Convergent validity	Discriminative validity	Comparison with gold standard	Comparison with other instruments	Comparison between subgroups
Brooks & Benson (2013) - SCQ							V	I	V	V		V			
Charman et al. (2007) - SCQ						I		A	V	V	V		V	V	
Corsello et al. (2007) - SCQ						V		V	V	V	V	V	V	V	
Ung et al. (2016) - SCQ						I	V	I	V	V	V	V	V	V	V
Mouti et al. (2019) - SCQ						V	V	V	V	V		V	V	V	V
Aldosari et al. (2019) - SCQ				A	A	V	V	V	V	V		V	V	V	V
Mesibov et al. (1989) - CARS						D		I	V						
Ooi et al. (2011) - CBCL						V		V	V						
Deb et al. (2009) - DBC-ASA							I	I	V	V					
Duda et al. (2016) - MARA						V	I	D	I	I					
Özdemir & Diken (2018) - Turkish AABC	V	V	V	V	V	V	V	V	A	V	A	V	V		A

V = very good, A = adequate, D = doubtful, I = inadequate

[1] Empty cell indicates no evidence or not applicable

Table 3 - Quality of studies on measurement properties for intellectual disabilities

STUDY & TOOL	Content validity[1]					Structural validity[1]	Internal consistency[1]	Cross-cultural validity[1]	Reliability[1]	Criterion validity[1]	Construct validity[1]		Responsiveness[1]		
	Asking patients			Asking experts											
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness						Convergent validity	Discriminative validity	Comparison with gold standard	Comparison with other instruments	Comparison between subgroups
Braatveit et al. (2018) – HASI	I	I	I	I	I		I	A	A	V	A	A	V		
Ford et al. (2008) - HASI	I	I	I	I	I		I	A	V	V	A	A	V		
Hayes (2002) - HASI	I	I	I	I	I		I	A	V	V	V	V	V	V	V
Sondenaa et al. (2007) - HASI	I	I	I	I	I	A	V	I	V	V	A		V		
Sondenaa et al. (2011) - HASI	I	I	I	I	I		V	I	V	V	A	A	V	A	
Sondenna et al. (2008) - HASI	I	I		I	I		I	A	V	V			V		
To et al. (2015) - HASI	I	I	I	I	I		I	I		V	V	V	V		
McKenzie et al. (2012a) - LDSQ	I	I	I	I	I		I	I			I	A			
McKenzie et al. (2015) - LDSQ	I	I	I	I	I		I	D			A		A		
McKenzie et al. (2014) - CAIDS-Q	I	I	I	I	I			A		A	D		A		
McKenzie et al. (2012b) - CAIDS-Q	I	I	I	I	I		V	I		I	A	V	A		

Table 3 - Quality of studies on measurement properties for intellectual disabilities

STUDY & TOOL	Content validity[1]					Structural validity[1]	Internal consistency[1]	Cross-cultural validity[1]	Reliability[1]	Criterion validity[1]	Construct validity[1]		Responsiveness[1]		
	Asking patients		Asking experts												
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness						Convergent validity	Discriminative validity	Comparison with gold standard	Comparison with other instruments	Comparison between subgroups
McKenzie et al. (2012c) - CAIDS-Q	I	I	I	D	D	V	V	V		V	V	V	V		
McKenzie et al. (2019) - CAIDS-Q	I	I	I	I	D	I			V	V	V	V			
Geijsen et al. (2016) - SCIL	I	I	I	D	D	V	V	A	D	V			V		A
Nijman et al. (2016) - SCIL	I	I	I	I	I	V	V	A	D	V			V		
Trivedi (1977) - SIT	I	I	I	I	I		I	I		V	V	V	V	V	V
Kunen et al. (1996) - SIT	I	I	I	I	I		I	I		V	V	V		V	V
Rotatori & Epstein (1978)- SIT	I	I	I	I	I		I	I							
Sawyer & Whitten (1972) - QT	I	I	I	I	I		I	I		V	A			D	

V = very good, A = adequate, D = doubtful, I = inadequate

[1] Empty cell indicates no evidence of not applicable

Table 4 - Quality of the TOOL development for autism

TOOL	TOOL design							Cognitive interview (CI) study[2]				TOTAL TOOL DEVELOPMENT
	General design requirements					Concept elicitation[1]	Total TOOL design	General design requirements	Comprehen-sibility	Comprehen-siveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the TOOL was developed	Clear context of use	TOOL developed in sample representing the target population			CI study performed in sample representing the target population				
Vrancic et al. (2002) - ADI-TSS	V	V	V	V	A	I	I	A	I	A	I	I
Morales-Hidalgo et al. (2017) - EDUTEA	V	V	V	V	V	I	I	V	I	V	I	I
Kraijer & de Bildt (2005) - PDD-MRS	V	V	V	V	V	I	I	V	I	D	I	I
Cortés et al. (2018) - EVTEA-DI (PDD-MRS)	V	V	V	V	V	A	A	V	A	D	D	D
Heinrich et al. (2018) - DiBAS-R	V	V	V	V	V	D	D					D
Allison et al. (2012) - AQ-10	V	V	V	V	V	D	D					D
Baron-Cohen et al. (2006) – AQ-10	V	V	V	V		N	N					V
Booth et al. (2013) – AQ-10	V	V	V	V			V					V
Cederberg et al. (2018) - ASSQ	V	V	V	V		V	I	I				I
Kopp & Gillberg (2011) –	V	V	V	V	V	V	V					V

Table 4 - Quality of the TOOL development for autism

TOOL	TOOL design							Cognitive interview (CI) study[2]			TOTAL TOOL DEVELOPMENT
	General design requirements					Concept elicitation[1]	Total TOOL design	General design requirements	Comprehen-sibility	Comprehen-siveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the TOOL was developed	Clear context of use	TOOL developed in sample representing the target population			CI study performed in sample representing the target population				
Berument et al. (1999) - SCQ ASSQ-REV	V	V	V	V	V	V	V					V
Brooks & Benson (2013) - SCQ	V	V	V	V	V	N	V					V
Charman et al. (2007) - SCQ	V	V	V	V	V		V					V
Corsello et al. (2007) - SCQ	V	V	V	V	V		V					V
Ung et al. (2016) - SCQ	V	V	V	V	V		V					V
Mouti et al. (2019) - SCQ	V	V	V	V	V	V	V					V
Aldosari et al. (2019) - SCQ	V	V	V	V	V	V	V					V

Table 4 - Quality of the TOOL development for autism

TOOL	TOOL design							Cognitive interview (CI) study[2]				TOTAL TOOL DEVELOPMENT
	General design requirements					Concept elicitation[1]	Total TOOL design	General design requirements	Comprehen-sibility	Comprehen-siveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the TOOL was developed	Clear context of use	TOOL developed in sample representing the target population			CI study performed in sample representing the target population				
Mesibov et al. (1989) - CARS	V	V	V	V	V		V					V
Ooi et al. (2011) – CBCL	V	V	V	V	V	N	V					V
Deb et al. (2009) - DBC-ASA	V	V	V	V	V	V	V	V	V	V	V	V
Duda et al. (2016) – MARA	V	V	V	V	V	I	I					I
Özdemir & Diken (2018) - Turkish AABC	V	V	V	V	V	V	V	V	A	A	A	A

V = very good, A = adequate, D = doubtful, I = inadequate

[1] When the TOOL was not developed in a sample representing the target population, the concept elicitation was not further rated

[2] Empty cells indicate that a CI study (or part of it) was not performed

Table 5 - Quality of the evidence for measurement properties of the autism screening tools

	ADI-TSS		**EDUTEA**		**PDD-MRS**		**DiBAS-R**		**AQ-10**		**ASSQ/ASSQ-REV**	
	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low
Content validity[1]	-	L	+	M	-	M	+	M	?	L	?	L
Relevance	+	H	+	M	+	M	+	M	+	M	?	L
Comprehensiveness	+	M	+	M	+	M	+	M	?	L	?	L
Comprehensibility	-	L	+	M	-	M	+	M	?	L	?	L
Structural validity	-	VL	+	M	+	M	+	M	?	VL	-	L
Internal consistency	-	VL	+	H	+	M	+	M	+	M	-	L
Cross-cultural validity	-	VL	-	M	+	M	-	L	?	VL	-	L
Reliability	-	L	+	H	-	M	+	M	-	L	-	L
Discriminative validity	+	M	+	M	+	M	+	M	+	M	+	M
Criterion validity	+	M	+	H	+	H	+	M	-	L	-	L
Construct validity	+	M	+	M	+	H	+	M	+	L	-	L
Comparison to gold standard tool	+	M	+	H	+	M	+	M	-	L	-	L

Table 5 - Quality of the evidence for measurement properties of the autism screening tools

	SCQ		**CARS**		**CBCL**		**DBC-ASA**		**MARA**		**AABC**	
	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low	OVERALL RATING + / - / ?	QUALITY OF EVIDENCE High, moderate, low, very low
Content validity[1]	+	M	?	L	+	M	+	M	+	M	+	M
Relevance	+	M	?	L	+	M	+	M	+	M	+	M
Comprehensiveness	+	M	?	L	+	M	+	M	+	M	+	M
Comprehensibility	-	M	?	L	+	M	+	M	+	M	+	M
Structural validity	+	M	+	M	+	M	+	M	+	M	+	H
Internal consistency	+	M	+	M	+	M	+	M	-	L	+	M
Cross-cultural validity	+	M	+	M	+	M	-	L	?	VL	-	L
Reliability	+	M	+	M	+	M	+	M	-	L	+	M
Discriminative validity	+	H	-	L	+	M	+	M	-	L	+	M
Criterion validity	+	H	-	L	+	M	+	M	-	L	+	M
Construct validity	+	M	-	VL	+	M	+	M	-	L	+	M
Comparison to gold standard tool	+	M	-	L	-	L	-	L	+	M	-	M

High = H, Moderate = M, Low = L, Very low = VL

Score: + = sufficient; - = insufficient; ? = indeterminate; ± = inconsistent

[1] These criteria refer to the construct, population, and context of use of interest in the systematic review.

Table 6 - Quality of the TOOL development for intellectual disabilities

TOOL	TOOL design							Cognitive interview (CI) study[2]				TOTAL TOOL DEVELOPMENT
	General design requirements					Concept elicitation[1]	Total TOOL design	General design requirements	Comprehensibili-ty	Compreh-eniveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the TOOL was developed	Clear context of use	TOOL developed in sample representi-ng the target population			CI study performed in sample representing the target population				
Braatveit et al. (2018) - HASI	V	V	V	V	V	I	I					I
Ford et al. (2008) - HASI	V	V	V	V	V	I	I					I
Hayes (2002) - HASI	V	A	V	V	A	I	I					I
Sondenaa et al. (2007) - HASI	V	V	V	V	V	D	D					D
Sondenaa et al. (2011) - HASI	V	V	V	V	V	D	D					D

Table 6 - Quality of the TOOL development for intellectual disabilities

TOOL	TOOL design							Cognitive interview (CI) study[2]			TOTAL TOOL DEVELOPMENT
	General design requirements					Concept elicitation[1]	Total TOOL design	General design requirements	Comprehensibili-ty	Compreh-eniveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the TOOL was developed	Clear context of use	TOOL developed in sample representi-ng the target population			CI study performed in sample representing the target population				
Sondenna et al. (2008) - HASI	V	V	V	V	V	I	I					I
To et al. (2015) - HASI	V	V	V	V	V	I	I					I
McKenzie et al. (2012a) - LDSQ	V	V	V	V	V	I	I					I
McKenzie et al. (2015) - LDSQ	V	V	V	V	V	I	I					I
McKenzie et al. (2014) - CAIDS-Q	V	V	V	V	V	I	I					I

Table 6 - Quality of the TOOL development for intellectual disabilities

TOOL	TOOL design							Cognitive interview (CI) study[2]				TOTAL TOOL DEVELOPM-ENT
	General design requirements					Concept elicitation[1]	Total TOOL design	General design requirements CI study performed in sample representing the target population	Comprehensibili-ty	Comprehen-siveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the TOOL was developed	Clear context of use	TOOL developed in sample representi-ng the target population							
McKenzie et al. (2012b) - CAIDS-Q	V	V	V	V	V	I	I	A	A		A	I
McKenzie et al. (2012c) - CAIDS-Q	V	V	V	V	V	A	A	V	V		V	A
McKenzie et al. (2019) - CAIDS-Q	V	V	V	V	V							V
Geijsen et al. (2016) - SCIL	V	V	V	V	V	D	D	V	V		V	D
Nijman et al. (2016) - SCIL	V	V	V	V	V	I	I					I

Table 6 - Quality of the TOOL development for intellectual disabilities

TOOL	TOOL design							Cognitive interview (CI) study[2]				TOTAL TOOL DEVELOPM-ENT
	General design requirements					Concept elicitation[1]	Total TOOL design	General design requirements CI study performed in sample representing the target population	Comprehensibili-ty	Comprehen-iveness	Total CI study	
	Clear construct	Clear origin of construct	Clear target population for which the TOOL was developed	Clear context of use	TOOL developed in sample representi-ng the target population							
Trivedi (1977) - SIT	V	A	A	V	D	I	I					I
Kunen et al. (1996) - SIT	V	V	V	V	V	D	D					D
Rotatori & Epstein (1978)- SIT	A	A	A	D	A	I	I					I
Sawyer & Whitten (1972) - QT	V	V	V	V	V	I	I					I

V = very good; A = adequate; D = doubtful; I = inadequate; NA = not applicable

[1] When the TOOL was not developed in a sample representing the target population, the concept elicitation was not further rated

[2] Empty cells indicate that a CI study (or part of it) was not performed

Table 7 - Quality of the evidence for measurement properties of the screening tools for intellectual

	HASI		SCIL		SIT		CAIDS-Q		LDSQ		QT	
	OVERALL RATING	QUALITY OF EVIDENCE	OVERALL RATING	QUALITY OF EVIDENCE	OVERALL RATING	QUALITY OF EVIDENCE	OVERALL RATING	QUALITY OF EVIDENCE	OVERALL RATING	QUALITY OF EVIDENCE	OVERALL RATING	QUALITY OF EVIDENCE
	+ / - / ?	High, moderate, low, very low	+ / - / ?	High, moderate, low, very low	+ / - / ?	High, moderate, low, very low	+ / - / ?	High, moderate, low, very low	+ / - / ?	High, moderate, low, very low	+ / - / ?	High, moderate, low, very low
Content validity[1]	-	L	+	M	?	VL	+	M	+	M	+	M
Relevance	+	M	-	M	?	VL	+	M	+	M	+	M
Comprehensiveness	-	L	+	H	?	VL	+	M	+	M	+	H
Comprehensibility	-	L	-	M	?	VL	+	M	+	M	+	M
Structural validity	?	VL	+	H	?	VL	+	M	-	L	-	L
Internal consistency	-	L	+	H	-	VL	+	M	-	L	-	L
Cross-cultural validity	+	M	+	H	-	VL	+	M	-	L	?	VL
Reliability	-	L	-	M	+	M	+	M	-	L	-	L
Discriminative validity	+	M	+	M	-	L	+	M	+	M	-	L
Criterion validity	+	H	+	H	±	M	+	M	+	M	+	M
Construct validity	+	H	-	M	±	M	+	H	+	M	+	M
Comparison to gold standard tool	+	M	+	H	-	L	+	M	+	M	-	L

High = H, Moderate = M, Low = L, Very low = VL

Score: + = sufficient; - = insufficient; ? = indeterminate; ± = *inconsistent*

[1] These criteria refer to the construct, population, and context of use of interest in the systema

Table 8 - Characteristics of the studies about screening tools for autism

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
1	Aldosari et al. (2019), Qatar	Validation of the Arabic version of the Social Communication Questionnaire (SCQ)	Between groups (ASD & Non-ASD)	SCQ: 40-items questionnaire	Childhood Autism Rating Scale (CARS), Autism Diagnostic Interview-Revised (ADI-R) & Autism Diagnostic Observation Schedule (ADOS)	Children: N = 412, ASD: N = 206, Non-ASD: N = 206, Total sample age range: 5 – 12yrs (M = 8.46, SD = 2.65)	5 – 10 mins	At cutoff of 15, sensitivity = .80, specificity = .97 At cutoff of 12, sensitivity = .89, specificity = .89 Youden's Index for cutoffs ranging from 11 to 15, sensitivity range was .90 - .80 and specificity range was .85 - .97
2	Allision et al. (2012), UK	Adapt the Autism Spectrum Quotient (AQ) (adult, adolescent & child versions) into short versions with good test accuracy (10 most discriminating items from the full versions to be included)	Between groups (ASD & Non-ASD)	Adult AQ-10, Adolescent AQ-10, Child AQ-10 and Q-CHAT-10 questionnaires	Participants had previous diagnosis	AQ-10 adult: N = 449, Control: N = 838, Age: >16yrs (M = 35.08, SD = 12.55) AQ-10 adolescent: N = 162, Control: N = 475, Range: 12 - 15yrs (M = 13.33, SD = 1.07) AQ-10 children & toddlers: N = 558, Control = 1,694, Range = 15 months – 11yrs	<10 mins	AQ-10 adult: sensitivity = .88, specificity = .91 PPV = .85, No NPV AQ-10 adolescent: sensitivity = .93, specificity = .95 PPV = .86, No NPV AQ-child: sensitivity = .95, specificity = .97 PPV = .94, No NPV AQ-toddler: sensitivity = .91, specificity = .89 PPV = .58, No NPV

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
3	Baron-Cohen et al. (2006), UK	1. The test the predictive ability of the revied AQ on adolescents 2. To determine whether Asperger Syndrome (AS) and classic autism yield similar results using the measure	Between groups (AS/HFA, ASD & Non-AS/ASD)	Revised AQ: 50-item questionnaire	None	Adolescents: N = 52, Age: >12yrs with AS/HFA (M = 13.6yrs, SD = 2.0), Range = 10.3 – 15.4 Adolescents: N = 79, With classic autism, (M = 12.5yrs, SD = 1.7), Range = 9.8 – 16.0 Control: N = 50, (M = 13.6yrs, SD = 1.8), Range = 10.1 – 16.5yrs	10 – 15 minutes	Sensitivity, specificity, PPV & NPV not reported
4	Berument et al. (1999), UK	To develop and test a screening questionnaire (Autism Screening Questionnaire – ASQ)	Between groups (PDD & Non-PDD)	ASQ: 40-items questionnaire (currently known as the SCQ)	Autism Diagnostic Interview-Revised (ADI-R)	Adults & children: N = 200, (M = 15.5yrs), Range = 4 – 40yrs	5 – 10 minutes	Sensitivity = .85, specificity = .75, PPV = .93, NPV = .55
5	Booth et al. (2013), UK	To evaluate the ability of the AQ-10 to discriminate between individuals with and without a confirmed ASD diagnosis	Between groups (ASD & Non-ASD)	AQ-10 adult questionnaire	None	ASD adults: N = 149, (M = 33yrs), Range = 17 – 75yrs Non-ASD adults: N = 134, (M = 29.6yrs), Range = 17 – 65yrs	5 minutes	AQ-10 adult: sensitivity = .80, specificity = .87, PPV & NPV not reported

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
6	Brooks & Benson (2013), USA	To assess the validity of the Social Communication Questionnaire in adults with a previous diagnosis of ID	Between groups (ASD+ID & ID only)	SCQ: 40-items questionnaire	Aberrant Behavior Checklist-Community (ABC-C) & Adaptive Behavior Assessment System-Second Edition (ABAS-II)	Overall: N = 69, Range = 18 – 40yrs, (M = 29.3, SD = 6.4) ASD & ID adults: N = 21 (M = 27.6yrs, SD = 5.8) ID only adults: N = 48 (M = 30.1yrs, SD = 6.5)	5 - 10 minutes	At cut-off score of 15 sensitivity = .71, specificity = .77, PPV = .58, NPV = .86 At a cut-off of 12 sensitivity = .86, specificity =0.60, PPV = .49, NPV = .91
7	Cederberg et al. (2018), USA	To examine whether the Autism Spectrum Screening Questionnaire (ASSQ) & Social Responsiveness Scale (SRS) can confirm ASD in high-ability youths with previous diagnosis	Cross-sectional (high-ability ASD)	1. SRS: 65-item parent and/or teacher report 2. ASSQ: 27-item parent and/or teacher report	Autism Diagnostic Observation Schedule-2 (ADOS-2) & ADI-R	Children: N = 23, (M = 11.93yrs, SD = 3.43), Range = 4.0 - 17.11yrs	SRS: 15 - 20 minutes ASSQ: 5 - 10 minutes	All participants in the study had a previous ASD diagnosis, no psychometrics reported.

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
8	Charman et al. (2007), UK	To compare the SCQ, SRS and the Children's Communication Checklist (CCC)	Cross-sectional (ASD)	1. SCQ: 40-item questionnaire 2. SRS: 65-item parent and/or teacher report 3. CCC: 70-rating scale	Autism Diagnostic Observation Schedule-Generic (ADOS-G), ADI-R & ICD-10	Children: N = 119, SCQ: (M = 10.2yrs, SD = 0.4), Range = 9.5 - 11yrs SRS: (M = 12.6yrs, SD = 0.4), Range = 11.8 – 13.2yrs CCC: (M = 12yrs, SD = 0.1), Range = 9.8 – 13.9yrs	SCQ: 5 - 10 minutes SRS: 15 - 20 minutes CCC: 5 – 15 minutes	1. SCQ - sensitivity = .86, specificity = .78, PPV = .74, NPV = .88 2. SRS - sensitivity = .78, specificity = .67, PPV = .63, NPV = .81 3. CCC - sensitivity = .93, specificity = .46, PPV = .56, NPV = .90
9	Corsello et al. (2007), UK	To investigate the clinical screening abilities of the SCQ in a large young American sample of children with & without ASD	Between groups (ASD & Non-ASD)	1. SCQ: 40-item questionnaire	Autism Diagnostic Observation Schedule (ADOS) & ADI-R	Children: N = 590 ASD = 439 Non-ASD = 151 Range = 2 - 16yrs	5 - 10 minutes	At cut-off of ≥15, Sensitivity = .71, specificity = .71, PPV = .88, NPV = .45 At cut-off of ≥12, Sensitivity = .82, specificity = .56, PPV = .84, NPV = .51

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
10	Cortés et al. (2018), Spain	To validate the Spanish adaptation of the Pervasive Developmental Disorder in Mentally Retarded Persons Scale (PDD-MRS)	Cross-sectional (ID)	Escala de Valoración del Trastorno del Espectro Autista en Discapacidad Intelectual (EVTEA-DI)	CARS	Adolescents and adults: N = 979 Age range = 18 - 70yrs, (M = 42.4, SD = 13.9)	Not available	At CARS cut-off of ≥30, sensitivity = .70, specificity = .91, PPV = .73, NPV = .90. At YI of 8, the EVTEA-DI sensitivity = .84, specificity = .83 At CARS cut-off of ≥25.5, sensitivity = .58, specificity = .56, PPV = .84, NPV = .51
11	Deb et al. (2009), UK	To examine the validity of the Developmental Behaviour Checklist-Autism Screening Algorithm (DBC-ASA) as a screening tool for autism in children with ID	Retrospective; Cross-sectional (ID only)	DBC-ASA: 29-item scale subset from the DBC	None	Children: N = 109, Age range: 3 - 17yrs	5 - 10 minutes	1. At cut-off score of 18 sensitivity = .92, specificity = .95, PPV = .87, NPV = .80 2. At a cut-off of 17 sensitivity = .95, specificity = .42. 3. A score of 20 however, provided a best fit with sensitivity = .90, specificity = .60
12	Duda et al. (2016), USA	To test the sensitivity and specificity of the Mobile Autism Risk Assessment (MARA) in a clinical sample of children referred for developmental/behavioural concerns	Cross-sectional (Developmental disorders including ASD)	MARA: electronically administered 7-item parent questionnaire ASD screener	ADOS	Children: N = 222, Age: 16 months - 17yrs, Median age = 5.8yrs	5 minutes	Sensitivity = .90, specificity = .80, PPV = .67, NPV = .95

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
13	Heinrich et al. (2018), Germany	1. To validate the diagnostic ability of the Diagnostic Behavioral Assessment for Autism Spectrum Disorder-Revised (DiBAS-R) in adults with ID 2. To assess the impact of the level of ID on the diagnostic accuracy of the tool	Cross-sectional, between group (ASD/ID & ID only)	DiBAS-R: 19-item care-giver screening scale	Autism-Checklist (ACL), ADOS, ADI-R, PDD-MRS & Music-based Scale for Autism Diagnostics (MUSAD)	Adults: N = 381, N = 289 (ID only), N = 92 (ASD/ID), (M = 40.5yrs, SD = 13.4), Range = 16 - 75yrs	5 minutes	1. Overall sensitivity = .82, specificity = .67, PPV = .44, NPV = .92 2. Mild-to-moderate ID: sensitivity = .79, specificity = .84, PPV = .51, NPV = .95 3. Severe-to-profound ID: sensitivity = .83, specificity = .34, PPV = .40, NPV = .79

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
14	Kopp & Gillberg (2011), Sweden	1. Present the extended, revised version of the Autism Spectrum Screening Questionnaire (ASSQ)-Revised Extended Version (ASSQ-REV) 2. Test the validity of the ASSQ-REV against the ASSQ 3. Examine its ability to discriminate between ASD and non-ASD cases 4. Analyse whether single items of the ASSQ-REV are more often endorsed in girls than in boys 5. Find the best predictors of ASD (vs non-ASD) in girls and boys	Between groups 1. ASD girls & ASD boys 2. ADHD girls & ADHD boys	ASSQ-REV: 45-item questionnaire	None	Children: N = 191, Age range: 6 - 16yrs	Not available	There was no significant difference between ASD girls and boys on the ASSQ-REV. The ASSQ seemed to identify more ASD than the ASSQ-REV in both groups (boys & girls). However, certain individual items from the ASSQ-REV seemed to discriminated boys with ASD from girls.

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
15	Kraijer & De Bildt (2005), Netherlands	1. To describe and discuss the Pervasive Developmental Disorder in Mentally Retarded Persons (PDD-MRS) Scale with focus on the four aspects of the construction of the scale. 2. Validating the scale against the ADOS	Cross-sectional (Profound ID)	PDD-MRS: 12-item questionnaire	ADOS	Child through adult: N = 1230, Age range: 2 - 80yrs	10 - 20 minutes	Sensitivity = .92, specificity = .92. No PPV & NPV reported
16	Mesibov et al. (1989), USA	To examine the suitability of the Childhood Autism Rating Scale (CARS) for diagnosing adolescents and adults with autism	Within group, longitudinal (15yrs)	CARS: 15-item scale	None	Out of over 1,500 children assessed over the 15yrs with the CARS, only N = 89 were diagnosed before age 10yrs (M = 8.7yrs) and again after age 13yrs (M = 15.9yrs)	5 - 10 minutes	At a cut-off score of 30, 81% retained their initial diagnoses for autism At a cut-off score of 27, 92% retained their initial diagnosis for autism

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
17	Morales-Hidalgo et al. (2017), Spain	1. To design and evaluate the psychometric properties of the EDUTEA: A DSM-5 teacher screening questionnaire for ASD & Social Communication Disorder (SCD) in school settings 2. Assessing the ability of the EDUTEA to discriminate between ASD/SCD and ADHD	Between groups (ASD/SCD, ADHD & No disorder)	EDUTEA: 11-item questionnaire	Childhood Asperger Syndrome Test (CAST), ADI-R, ADOS-2, CBCL & Schedule for Affective Disorders and Schizophrenia (K-SADS-PL)	Children: N = 291, N = 175 (ASD/SCD+ADHD), N = 116 (control) Age range: 3 - 12yrs	Not available	At a proposed cut-off of 10 1. Psychometric properties; sensitivity = .87, specificity = .91, PPV = .86, NPV = .99 2. Ability to discriminate between ASD/SCD and ADHD; sensitivity = .83, specificity = .73
18	Mouti et al. (2019), Australia	1. To differentiate between ASD & ADHD using the SCQ & determine a clinical cutoff score for doing so 2. Will children with ADHD score higher than those without in all SCQ domains?	Between groups (ASD & ADHD & No disorder)	SCQ	Participants had previous diagnosis	Children & adolescents: N = 162, Age range: 6 - 17yrs (M = 11.27, SD = 3.28)	5 – 10 minutes	SCQ total: at a proposed cut-off of 13, ASD vs ADHD sensitivity = .96, specificity = .87 SCQ social communication: at a cutoff of 7, ASD vs ADHD sensitivity = .96, specificity = .73 SCQ repetitive behaviours: at a cutoff of 3, ASD vs ADHD sensitivity = .81, specificity = .74

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
19	Ooi et al. (2011), Singapore	1. To test the ability of each of the eight Child Behavior Checklist (CBCL) syndromes to differentiate the ASD group from the other groups 2. To test which CBCL syndromes significantly differentiate the ASD group from the comparison group when all syndromes are used as predictors in a single analysis 3. To test which CBCL items differentiate the ASD group from the comparison groups 4. To derive and test an ASD scale comprised of items with significant differentiating ability	Between groups (ASD, ADHD, Undiagnosed and No disorder)	CBCL: 118-items parent-rated form	None	Children: N = 1,265, N = 86 (ASD), N = 543 (ADHD), N = 200 (undiagnosed), N = 436 (control), (M = 9.06yrs, SD = 2.45), Range: 4 - 18yrs	30 minutes to 1 hour	1. The ability of the CBCL syndromes to differentiate between ASD and comparison group sensitivity ranged from .50 – .78, specificity range .59 – .87 2. 9 specific items from the CBCL indicated to be predictive of ASD and formed the basis of the construction of 9-item ASD scale with sensitivity ranging from .68 – .78, specificity range .73 – .92

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
20	Özdemir & Diken. (2019), Turkey	1. To develop an assessment tool for ASD in Turkey based on the Autism Behavior Checklist (ABC) 2. To validate the adapted instrument, the AABC (Adapted Autism Behaviour Checklist)	Cross-sectional between groups (ASD & ID)	AABC (Adapted Autism Behaviour Checklist), 57-item questionnaire	Gilliam Autism Rating Scale-2 Turkish Version (GARS-2 TV)	Children & adolescents: N = 1133. N = 969 (ASD), N = 164 (ID), Range: 3 – 15yrs, (M = 9yrs, SD = 3.43)	Not available	At a cut-off of 13 sensitivity = .87, specificity = .82
21	Ung et al. (2016), USA	1. To replicate the internal consistency and convergent validity of the SCQ 2. To compare the sensitivity & specificity at different cut-off scores optimal for distinguishing between ASD and Non-ASD	Between groups (ASD and Non-ASD)	SCQ: 40-items questionnaire	CARS-2 and the Vineland Adaptive Behavior Scale (VABS)	Children: N = 76, N = 33 (ASD), N = 43 (non-ASD), Range: 4 – 12yrs, (M = 6.70yrs, SD = 1.86)	5 – 10 minutes	1. At cut-off ≤11 sensitivity = .82, specificity = .37, PPV = .50, NPV = .73 2. At cut-off ≤15 sensitivity = .70, specificity = .67, PPV = .62, NPV = .74 3. At cut-off ≤22 sensitivity = .30, specificity = .91, PPV = .71, NPV = .63

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
22	Vrancic et al. (2002), Argentina	To develop and validate a screening questionnaire for autism administered over the telephone; Autism Diagnostic Inventory-Telephone Screening in Spanish (ADI-TSS)	Between groups (ASD & Non-ASD)	ADI-TSS: 47 questions designed using the ADI-R	ADI-R	Children: N = 59, N = 30 (ASD), N = 29 (non-ASD), Age range: 5 - 30yrs	20 - 40 minutes	Overall sensitivity = 1, specificity = .66

* Sensitivity is described as the ability of a screening tool correctly identify all persons who have (true positive) the condition of interest while specificity is the tool's ability to identify all persons who do not have (true negative) the condition of interest. Also, the positive predictive value (PPV) measures the probability that persons with a positive screening result truly have the condition, while the negative predictive value (NPV) is the probability that those with a negative screening result truly do not have the condition.

Table 9 - Characteristics of the studies about screening tools for intellectual disabilities

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
1	Braatveit et al. (2018), Norway	To validate the Hayes Ability Screening Index (HASI) in persons with substance use disorder	Cross-sectional (ID)	HASI: 4 subtests	WAIS-IV	Adults: N = 84, (M = 33.31yrs, SD = 11.65), Range: 19 – 64yrs	10 – 15 minutes	At a cut-off of 85, sensitivity = 1, specificity = .65 At a cut-off of 80.7, sensitivity = 1, specificity = .81 No PPV & NPV
2	Ford et al. (2008), UK	To evaluate the effectiveness of the HASI in identifying LD in adolescent offenders	Cross-sectional (young offenders with ID)	HASI: 4 subtests	WISC-IV, WAIS-III	Adolescents: N = 71, (M = 16yrs, SD = 2), Range: 10 – 19yrs	10 – 15 minutes	At a cut-off of 80.2, sensitivity = .8, specificity = .65 No PPV & NPV
3	Geijsen et al. (2016), Norway	1. To examine the predictive validity of the Screener for Intelligence & Learning Disabilities (SCIL) 2. To estimate the prevalence of ID among police suspects	Cross-sectional (suspects in police custody with ID)	SCIL: 14-items	WAIS-III-NL	Adults: N = 178, (M = 31.7rs, SD = 11.2), Range: 18 – 63yrs	10 minutes	1. Sensitivity = .72, specificity = .71 No PPV & NPV 2. Prevalence rate was 39.3%

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
4	Hayes (2002), Australia	Report on the effectiveness of the HASI	Cross-sectional, between groups (ID & No-ID)	HASI: 4 subtests	K-BIT, WASI-R, WISC-III	Youth & Adults: N = 567, ≤18yrs (N= 161), ≥18yrs (N = 406)	5 – 10 minutes	Adults: at cut-off of 85, sensitivity = .82, specificity = .72 Youth cut-off is recommended as 90 (no sensitivity or specificity is given)
5	Kunen et al. (1996), USA	1. Determining the degree of correlation between the Slosson Intelligence Test-Revised (Slosson-R) and the Stanford-Binet Intelligence Scale (Stanford-Binet) 2. Determine the degree of consistency of categorisation about intellectual functioning of individuals between the Slosson-R and the Stanford-Binet	Within-subject (ID)	Slosson-R: 187-items	Stanford-Binet Intelligence Scale	Children & Adults: N = 191, Range: 5 - 69yrs, (M = 27, SD = 12.69)	15 – 20 minutes	Results suggest good concurrent validity of the Slosson-R 1. Slosson-R mean IQ was not significantly different from the Stanford-Binet 2. For broad categorisation, the Slosson-R and Stanford-Binet have good match rate; 95.1% for mental retardation and 92.3% for no mental retardation 3. Correlation between the two instruments was .92

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
6	McKenzie et al. (2012a), UK	To assess the validity of the Learning Disability Screening Questionnaire (LDSQ) in forensic settings using an independent measure of ID based on diagnostic criteria	Between groups (ID & No-ID)	LDSQ: 7-items	Specific FSIQ not indicated	Adults: N = 94 ID group: N = 62, Range: 18 - 61yrs (M = 36yrs 7mo, SD = 11yrs 6mo) Non-ID group: N =32, Range: 22 - 62yrs (M = 40yrs, SD = 16yrs)	5 minutes	Sensitivity = .82, specificity = .88, PPV = .93, NPV = .74
7	McKenzie et al. (2014), UK	To assess the performance of the seven-subset short form of the Wechsler Adult Intelligence Scale for Children-fourth edition (WISC-IV) and the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q) in correctly identifying individuals with an ID diagnosis	Between groups (ID & No-ID)	CAIDS-Q: 7-item	WISC-IV: 7-item	Children and adolescents: N = 276 (ID = 106), Mean age = 135.7 months, SD = 36.9; (No-ID = 170), Mean age = 131.7 months, SD = 39.4	CAIDS-Q: 5 minutes WISC-IV short form: Not available	WISC-IV sensitivity = .92, specificity = .91 CAIDS-Q sensitivity = .89, specificity = .88

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
8	McKenzie et al. (2012b), UK	To validate the CAIDS-Q for use within the criminal justice system for young people with ID	Between groups (ID & No-ID)	CAIDS-Q: 7-item	WISC-IV, WAIS-III	Adolescents: N = 23 ID group: N = 8, Range: 15 – 18.16yrs (M = 16.89yrs, SD = 0.87) Non-ID group: N = 15, Range: 10.8 – 17.16yrs (M = 14.77yrs, SD = 1.75)	5 minutes	PPV = 1, NPV = 1
9	McKenzie et al. (2012c), UK	To evaluate some of the psychometric properties of the screening tool for ID, the CAIDS-Q with two age groups	Between groups (ID & No-ID)	CAIDS-Q: 7-item	WISC-IV, WAIS-III	Children: N = 130 (ID = 61), Range: 8 – 11.92yrs (M = 10yrs, SD = 1.1); (No-ID = 69), Range: 8 – 11.92yrs (M = 9.73yrs, SD = 1.15) Adolescents: N = 156 (ID = 77), Range: 12 – 17.1yrs (M = 14.38yrs, SD = 1.33); (No ID = 79), Range: 12 – 18yrs (M = 14.17yrs, SD = 1.35)	5 minutes	Children cut-off 62: sensitivity = .97, specificity = .86 Adolescent cut-off 64: sensitivity = .96, specificity = .85

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
10	McKenzie et al. (2015), Scotland	Validating the LDSQ against the Weschler Adult Intelligence Scale, Fourth edition (WAIS-IV)	Between groups (ID & No-ID)	LDSQ: 7-items	WAIS-IV	Adults: N = 83 (ID = 59), Range: 16 – 66yrs, (M = 32.1yrs, SD = 15.4); (No ID = 24), Range: 16 – 55yrs, (M = 30.8yrs, SD = 11.9)	5 minutes	Sensitivity = .92, specificity = .92
11	McKenzie et al. (2019), Scotland	Evaluating the psychometric properties of the CAIDS-Q in a paediatric neurodevelopmental setting	Cross-sectional between groups (ID & No-ID)	CAIDS-Q: 7-items	WISC-IV, ABAS II/III	Children & Adolescents: N = 181 Age range: 6 – 18yrs, ID: N = 54, Age range: 72 – 199 months (M = 117, SD = 29.9); No ID = 127, Age range: 72 – 210 months (M = 120.1, SD = 32.7)	5 minutes	Sensitivity = 1, specificity = .88, PPV = 1, NPV = .78
12	Nijman et al. (2018), Netherlands	To develop and test a time-efficient screener for intelligence and learning disabilities (SCIL) in the Dutch language for mild to borderline intellectual disabilities (MBID)	Cross-sectional Between groups (Mild to Borderline ID - MBID) & No-MBID	SCIL: 14-items	WAIS, WISC	Study 1: Adults: N = 318, Mean age = 31.5yrs (SD = 13.1yrs) Study 2: Adolescents: N = 305, Range 12 – 17yrs (M = 14.5yrs, SD = 1.6)	10 – 15 minutes	Study 1 cut-off score 19: Sensitivity = .83, specificity = .89 Study 2: varied based on the age range 12 - 13yrs not recommended 14 - 15yrs cut-off score 16; Sensitivity = .85, specificity = .82 16 - 17yrs cut-off score 18; Sensitivity = .80, specificity = .84

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
13	Rotatori & Epstein (1978), USA	1. To assess the test-retest reliability of the Slosson Intelligence Test (SIT) 2. To assess the ability of personnel with limited psychometric knowledge and experience to administer the test	Within-subject (Profound & Severe ID)	SIT: 187-items	None	Children and adolescents: N = 53, Range = 5yr 3mo - 16yr 9mo (M = 12.72, SD = 3.54)	10 - 20 minutes	1. Test-retest coefficients ranged between .91 and .96, indicating that the teachers could administer the test 2. Authors view the SIT to be a reliable screener
14	Sawyer & Whitten (1972), USA	To investigate the concurrent validity of the Quick Test (QT) in the mildly mentally retarded	Within-subject (Mild ID)	QT: 3 individual forms, 50-items	WISC	Children and adolescents: N = 27, Mean age = 12.1yrs	3 - 10 minutes	Correlation ranged from .08 - .99. Consistent correlation with the Picture Arrangement subtest on the WISC
15	Søndenaa et al. (2007), Norway	Validation of the Norwegian Hayes Ability Screening Index (HASI) against the WAIS-III	Within-subject (Referred for neuropsychological test)	HASI: 4 subtests	WAIS-III	Adults: N = 73, Range: 17 – 60yrs, (M = 33.3yr, SD = 12.5)	10 – 15 minutes	At the stated cut-off of 85, sensitivity = 1, specificity = .57. Alternate cut-off of 81, sensitivity = .95, specificity = .72 *r* = .81
16	Søndenaa et al. (2011), Norway	Validation of the Hayes Ability Screening Index (HASI) against the WASI	Cross-sectional Within-subject (psychiatric patients)	HASI: 4 subtests	WASI	Adults: N = 50, Range: 18 – 72yrs (M = 41.9, SD = 13.6)	10 – 15 minutes	At cut-off of 85, sensitivity = 1, specificity = .35, *r* = .67

S/No	Author (Date) & Country of Origin	Aim	Design	Measure	Validation Tool	Participants	Time to Administer	Outcome
17	Søndenaa et al. (2008), Norway	1. To estimate prevalence & nature of ID in Norwegian inmates. 2. Validation of the Norwegian Hayes Ability Screening Index (HASI) against the WASI	Cross-sectional Within-subject (ID)	HASI: 4 subtests	WASI (Norwegian version)	Adults: N = 143, Range: 19 – 68yrs (M = 34.6)	10 – 15 minutes	At cut-off of 85, sensitivity = .93, specificity = .72 At cut-off of 80, sensitivity = .87, specificity = .85, *r* = .72
18	To et al. (2015), Belgium	1. The validity of the Dutch HASI against the Dutch WAIS-III for substance abusers 2. To examine the effects of psychiatric disorder on the HASI results	Within-subject (Substance users with ID)	HASI: 4 subtests	WASI-III (Dutch version)	Adults: N = 90, Mean age = 32yrs, (SD = 9.795)	10 – 15 minutes	HASI/WAIS-III FSIQ, *r* = 0.694 At cut-off score of 85, sensitivity = .91, specificity = .80
19	Trivedi (1977), USA	To determine the reliability among the Wechsler Intelligence Scale for Children (WISC), SIT and Peabody Picture Vocabulary Test (PPVT)	Within-subject (ID)	SIT: 187-items PPVT: 175 stimuli words and pictures	WISC	Adolescents: N = 36, Range: 13 – 16yrs (M = 15.29yrs, SD = 0.64)	SIT: 10 -20 minutes PPVT: 20 - 30 minutes WISC: 45 - 65 minutes	Correlations were: WISC & PPVT range *r* = .60 – .76, WISC & SIT range *r* = .85 – .89, PPVT & SIT range *r* = .49 – .76

* Sensitivity is described as the ability of a screening tool correctly identify all persons who have (true positive) the condition of interest while specificity is the tool's ability to identify all persons who do not have (true negative) the condition of interest. Also, the positive predictive value (PPV) measures the probability that persons with a positive screening result truly have the condition, while the negative predictive value (NPV) is the probability that those with a negative screening result truly do not have the condition

4.3 Results

4.3.1 Search Results

Forty-one papers met the eligibility criteria (Figure 1), and 22 of these were about screening tools for autism, while 19 focused upon screening tools for intellectual disabilities. The quality ratings for the included studies are found in Tables 2 and 3. Additionally, sensitivity (true positive rate), specificity (true negative rate), positive predictive value or precision (the probability of screening positive and being correct), and negative predictive values (the probability of screening negative and being correct) for the tools were extracted and are reported in Tables 8 and 9.

4.4 Description and Characteristics of Studies

4.4.1 Autism Spectrum Disorder

There were a total of 12,240 participants across the 22 autism studies with age ranging from 15-months to 80-years. Studies with children younger than 11-years or older than 27-years old were only included if most of their participants were within the specified range of the inclusion criteria. Of the 12,240 participants, a little over 9,000 involved proxy respondents such parents, teachers, or caregivers of people with autism. Most of the studies were conducted in the United Kingdom (UK) (n = 7) with others spread across various countries, including the United States (US) (n = 5), Spain (n = 2) and one each from Germany, Sweden, Netherlands, Qatar, Australia, Turkey, Singapore and Argentina (Table 8). A variety of screening tools were used across the studies, including the Autism Screening Quotient (AQ-10) adolescent and adult versions (n = 3), Autism Screening Quotient (AQ-50) (n = 1), Autism Spectrum Screening Questionnaire (ASSQ) (n = 1), Autism Spectrum Screening Questionnaire-Revised Extended Version (ASSQ-REV) (n = 1), Social Communication Questionnaire (SCQ) (n = 7), Developmental Behavior Checklist-Autism Screening Algorithm (DBC-ASA) (n = 1), Childhood Autism Rating Scale (CARS) (n = 1), Mobile Autism Risk Assessment (MARA) (n= 1), Pervasive Developmental Disorder in Mentally Retarded Persons Scale (PDD-MRS) (n = 2), EDUTEA (n = 1), Child Behavior Checklist (CBCL) (n = 1), Adapted Autism Behaviour Checklist (AABC) (n = 1), and Autism Diagnostic Inventory-Telephone Screening in Spanish (ADI-TSS) (n =1).

Fifteen studies were between-group designs, one within-group design, and six single group designs. One study was longitudinal and included data collected over 15-years. Across the included 22 studies, two broad aims were discerned: (a) designing a short screening tool, and (b) validating the discriminative ability of tools. Those that focused on designing short tools were further categorised in two ways: (a) adapting existing tools into shorter versions, or (b) the development of entirely new tools. Eighteen out of 22 (approximately 82%) of the papers reviewed based their studies on existing tools developed over ten years ago. The remaining four (18%) considered tools that were developed in the last two to three years. The existing tools were mainly used with children, while the other studies reviewed focused upon adapting the tools for adolescents and adults.

4.4.2 Autism Screening Tools

4.4.2.1 The Autism Spectrum Quotient (AQ)

The Autism Spectrum Quotient (Baron-Cohen et al. 2001) is a short, easy to use and score, self-administered screener for adults with Asperger Syndrome or High Functioning Autism. It is comprised of 50 questions divided into five subsets of 10 questions each covering five domains – social skills, attention to detail, attention switching, communication and imagination. Over time, the AQ was adapted and modified to include adolescents (Baron-Cohen et al., 2006) while maintaining the original 50-item format. The AQ-50 child, AQ-50 adolescent and AQ-50 adult measures were adapted to create shorter versions by selecting the ten most discriminating items from each and validating the short tool (Allison et al., 2012). The AQ in different variations was used in three different studies (Allison et al., 2012; Baron-Cohen et al., 2006; Booth et al., 2013). The short version of the adolescent tool, the AQ-10 (Allison et al., 2012) had a sensitivity of .93, a specificity of .95 and a positive predictive value (PPV) of .86 while the AQ-50 (Baron-Cohen et al., 2006) tested with adolescents had a sensitivity of .89 and specificity of 1. Baron-Cohen et al. (2006) reported no PPV but commented that future research should explore this.

For adult participants (includes participants older than 18 years and/or 16 years of age in some instances) that employed the short AQ-10, Allison et al. (2012) found a sensitivity of .88 and specificity of .91 while Booth et al. (2013) found a sensitivity of .80 with a specificity of .87. All

three studies (Allison et al., 2012; Baron-Cohen et al., 2006; Booth et al., 2013) included participants with a previous diagnosis of autism.

While all the three studies that employed the AQ defined the constructs to be measured, the quality of evidence was rated as low. Specifically, content validity was rated as low since participant and expert involvement in the studies was unclear. Structural validity, internal consistency, reliability, construct validity, cross-cultural validity and criterion validity were all examined by Allison et al. (2012). Baron-Cohen et al. (2006) examined internal consistency and reliability with moderate evidence for cross-cultural validity. Booth et al. (2013) provided evidence for structural validity, while reliability and cross-cultural validity were undetermined. As such, the evidence for reliability was rated as low and the overall rating for cross-cultural validity was found to be low. To ensure that the psychometric properties of the AQ-10 were accurately captured, the authors were contacted for the manual, who responded that the tests and 'manuals' were those on the authors' website. In summary, although the psychometric results met the criteria for good tools (Glascoe, 2005) following the COSMIN guidelines, where the lowest score counts, the overall quality of evidence for the tool was rated as low.

4.4.2.2 Autism Spectrum Screening Questionnaire (ASSQ) and the Autism Spectrum Screening Questionnaire-Revised (ASSQ-REV)

Preliminary development of the ASSQ took place in Sweden for use within a prevalence study for high-functioning autism and Asperger syndrome in mainstream schools (Ehlers and Gillberg, 1993). The ASSQ is a 27-item checklist that can be completed by laypersons such as teachers or parents and was developed further in later studies (Ehlers et al., 1999). An extended version of the ASSQ-REV was developed for the early identification of girls with autism (Kopp & Gillberg, 2011). The original Swedish version of the ASSQ has been translated into multiple languages – Mandarin Chinese (Guo et al., 2011), English (Ehlers & Gillberg, 1993), Norwegian (Posserud et al., 2006), Finnish (Mattila et al., 2009), and Lithuanian (Lesinskiene, 2000).

Cederberg et al. (2018) examined the diagnostic accuracy of the ASSQs in adolescents previously diagnosed with high functioning autism. While participant gender and the psychometric properties of the measure were not reported, the authors reported that the ASSQ appeared sensitive to correctly identifying autism. Kopp & Gillberg (2011) examined the validity and accuracy of individual items for detecting autism in girls and boys aged 6 – 16yrs. Different

items showed considerable discriminative ability (AUC > .70, see Kopp & Gillberg, 2011) for those with autism versus typically developing children across genders. Both studies used participants who had a previous diagnosis of Autism. Like Cederberg et al. (2018), Kopp & Gillberg (2011) reported no sensitivity, specificity, PPV or negative predictive value (NPV).

Although the ASSQ was originally in Swedish, and has been translated into different languages, cross-cultural validity was rated as low using COSMIN due to insufficient evidence of its effectiveness in different cultures. Criterion validity, construct validity, internal consistency and reliability were rated as insufficient based on the combined evidence from both studies (Kopp & Gillberg, 2011; Cederberg et al., 2018). Neither of the two studies examined the content nor structural validity of the ASSQ. To ensure that all relevant evidence and information on the tool's development were examined, efforts were made to access the manual but were unsuccessful. No other studies utilising the ASSQ outside the West were found. The overall quality of the tool was rated as very low.

4.4.2.3 Social Communication Questionnaire (SCQ)

The SCQ formerly known as the Autism Screening Questionnaire (ASQ; Berument et al., 1999), was initially designed as a companion screening tool for the Autism Diagnostic Interview (ADI) (Snow, 2013). The SCQ is a brief 40-item parent or caregiver-report screening measure modelled after the ADI-R and has been used widely in research (Berument et al., 1999; Rutter et al., 2003). The measure has two versions; the lifetime version and the current version, both focusing on symptoms of autism most likely to be observed by the individual's principal caregiver. The caregiver must be familiar with the individual's developmental history and current behaviour. The SCQ is a screening tool and cannot be used for the diagnosis of autism. The measure is used for anyone 4-years old and above. The design allows for the comparison of symptoms across different groups of individuals such as children with language delays and those with medical conditions co-existing with autism. The SCQ is currently available in seventeen languages (Danish, Dutch, English, Finnish, French, German, Greek, Hebrew, Hungarian, Italian, Japanese, Korean, Norwegian, Romanian, Russian, Spanish and Swedish) and is used widely in research.

Seven studies (Aldosari et al., 2019; Brooks & Benson, 2013; Berument et al., 1999; Charman et al., 2007; Corsello et al., 2007; Mouti et al., 2019; Ung et al., 2016) utilised the SCQ. Five studies (Aldosari et al., 2019; Charman et al., 2007; Corsello et al., 2007; Mouti et al., 2019; Ung

et al., 2016) included samples of adolescents, one included adults with intellectual disabilities (Brooks & Benson, 2013), while one (Berument et al., 1999) was a development study and included children, teenagers and adults (age range: 4 – 40 years). Berument and colleagues (1999) recommended an optimal cut-off of 15 for differentiating those with and without autism. Using this cut-off, they reported a sensitivity of .85, specificity of .75, PPV .93 and NPV .55. In the other studies, the cut-off was varied to generate optimal values, depending on the age of the participants. For instance, Brooks & Benson (2013) using a cut-off of 15, reported that the sensitivity was .71, specificity .77, PPV .58 and NPV .86. However, when the cut-off was lowered to 12, the sensitivity was .86, specificity .60, PPV .49 and NPV .91. Similarly, Corsello et al. (2007) reported finding sensitivity of .71, specificity = .71, PPV = .88, and NPV = .45 at a cut-off of 15 while at a cut-off of 12 sensitivity was .82, specificity = .56, PPV = .84, and NPV = .51. However, as is typical with screening tools, lower cut-off scores will improve sensitivity, but at the expense of specificity.

Recently, Mouti et al. (2019) examined the optimal cut-off for differentiating between ASD, attention deficit and hyperactivity disorders (ADHD) and typically developing individuals. Their result showed that at a cut-off of score of 9, the SCQ showed excellent discriminative ability between ASD and Non-ASD with a sensitivity of .1 and specificity of .84. Additionally, Mouti et al. (2019) showed that at the cut-off of 13, ASD was clearly discriminated in individuals who were diagnosed as ASD only (sensitivity = .96, specificity = .87) or a combination of both ASD and ADHD (sensitivity = .87, specificity = .85). In the Arabic validation study, Aldosari et al. (2019) reported sensitivity and specificity of .80 and .97 respectively at the recommended cut-off score of 15. However, for a cut-off range between 11 and 15, the sensitivity varied between .90 and .80 while specificity varied between .85 and .97. Aldosari et al. (2019) also reported internal consistency of $\alpha = .92$.

Apart from Ung et al. (2016), who validated the SCQ against the Childhood Autism Rating Scale (CARS-2) only, all the other studies validated the SCQ against either the Autism Diagnostic Observation Schedule-2 (ADOS-2) or Autism Diagnostic Interview-Revised (ADI-R) or a combination of the CARS, ADOS-2 and ADI-R. Overall, the psychometric properties of the SCQ met the guidelines (Glascoe, 2005) for good tools, and the SCQ correlated well with the ADI-R (Berument et al., 1999).

Out of the seven studies reviewed, four (Berument et al., 1999; Corsello et al., 2007; Mouti et al., 2019; Aldosari et al., 2019) examined the structural validity with sufficient outcomes reported. Criterion validity and reliability were rated as excellent across all seven studies. All seven studies had clear constructs with five (Charman et al., 2007; Corsello et al., 2007; Ung et al., 2016; Mouti et al., 2019; Aldosari et al., 2019) providing sufficient evidence for the construct validity. There was an excellent outcome on the criterion validity across all seven studies. Five studies (Berument et al., 1999; Corsello et al., 2007; Mouti et al., 2019; Aldosari et al., 2019; Charman et al., 2007) rated positive had sufficient evidence for cross-cultural validity while the remaining two (Brooks & Benson, 2013; Ung et al., 2016) were rated negative with insufficient evidence. Soto et al. (2015) in their review of culturally adapted tools, reported that the Chinese validation study (Gau et al., 2011) of the SCQ had good test-retest reliability (r_{ICC} = .77 – .78) and internal consistency (α = .73 – .91). The authors (Gau et al., 2011) reported excellent concurrent validity ($r \leq .65$). Given that the SCQ is available in 17 languages, has been used across countries including Africa (Bozalek, 2013), across ethnicities, genders, ages, and widely employed in research, it meets several of the qualities for good cross-cultural validity as defined by COSMIN. The SCQ was rated overall as medium based on the evidence from the seven studies reviewed (Brooks & Benson, 2013; Berument et al., 1999; Charman et al., 2007; Corsello et al., 2007; Ung et al., 2016; Mouti et al., 2019; Aldosari et al., 2019) and previous work done by McConachie et al. (2015). Given the above results, the SCQ seems an appropriate tool to be considered for use within African nations, especially as very little training is required to score it.

4.4.2.4 Childhood Autism Rating Scale (CARS)

The CARS was developed by Schopler et al. (1980) as a diagnostic tool for children with autism. However, this measure, while meant to be diagnostic, was included because Mesibov et al. (1989) used it as a screening tool with adolescent participants, suggesting the CARS' potential utility as a screening instrument for autism. Nevertheless, Mesibov et al. (1989) did comment that the CARS was meant to be used as a diagnostic tool. The CARS is a 15-item rating scale that assesses behaviours associated with autism. The measure is meant to ease the identification of children with autism for parents, educators, clinicians, and other health care providers. The scale is available in English, Brazilian Portuguese, Lebanese, Japanese, Swedish, and French.

The second edition now includes a scale for identifying high functioning autism and a parent information form. Some training is required to administer the tool.

Although the CARS was initially validated for use with children, Mesibov et al. (1989), in their longitudinal study, examined its suitability for use with adolescents and adults with autism. Fifty-nine participants with a previous autism diagnosis were re-assessed, and the results showed that 81% (n = 48) retained their diagnosis. In comparison, 19% (n =11) of them received a revised diagnosis of no autism based on a cut-off score of 30. However, moving the score to a cut-off of 27 (to account for the mean difference in scores between the younger and older sample), 92% (n = 54) were accurately diagnosed. As a result of the improved diagnostic outcomes, Mesibov and colleagues recommended 27 as the cut-off for persons over the age of 13-years.

Based on COSMIN guidelines, content validity, internal consistency, and construct validity were rated as not determined, since it was unclear from the study whether these were tested. There was insufficient evidence for structural validity, and criterion validity. Cross-cultural validity was rated as positive with moderate evidence due to the availability of the measure in different translations. The evidence for reliability was moderate; however, this was based on the evidence from the only study found (Mesibov et al., 1989). Authors were contacted for more information on the development of the tool or for access to the relevant portion of the manual, unsuccessfully. A search was done to find other studies that reported the development of the measure or studies in which the CARS was used. One such study was identified (Schopler et al., 1980) which reported an internal consistency coefficient of $\alpha = .94$ and interrater reliability of .71. Two other studies (DiLalla & Rogers, 1994; Breidbord & Croudace, 2013) were also identified: DiLalla & Rogers (1994) presented the results of an exploratory factor analysis of the CARS while Breidbord & Croudace (2013) examined the interrater reliability and internal consistency from various studies. Based on the results of these studies (Schopler et al., 1980; DiLalla & Rogers, 1994; Breidbord & Croudace, 2013) and evidence from McConachie et al. (2015), internal consistency, structural validity and reliability were rated as moderate. The overall rating for the measure was medium based on COSMIN guidelines.

Additionally, as per the publisher's guidance, some training and specific educational qualification are required before using the CARS. Thus, it seems inappropriate for further consideration for screening adolescents in Africa.

4.4.2.5 Developmental Behavior Checklist-Autism Screening Algorithm (DBC-ASA)

The DBC-ASA (Brereton et al., 2002) is a 29-item autism screening measure derived from the Developmental Behavior Checklist (DBC). The DBC was revised and updated to the DBC2 in 2018. The parent version of the DBC is available in the following languages: Chinese, Arabic, Croatian, Dutch, French, Finnish, German, Greek, Hindi, Norwegian, Portuguese (Brazilian), Italian, Japanese, Spanish, Swedish, Turkish and Vietnamese.

Deb et al. (2009), screened a total of 109 children aged 3 – 17 years with intellectual disabilities for autism using the instrument. Forty-four of the children were between 3 – 9 years old, 50 of them between 10 – 15 years old and 15 participants were older than 15 years. A cut-off score of 19 for the 3 – 9 years olds yielded a sensitivity of 1 and specificity of .71 while a cut-off of 26 for the 10 – 15-year-olds yielded a sensitivity of .70 and specificity of .75. When a total population cut-off score of 20 was applied, sensitivity was .90 and specificity .60. The figures generated by Deb et al. (2009) differ from those obtained in Brereton et al. (2002) where a cut-off score of 14 yielded sensitivity of .86 and specificity of .55 and a cut-off score of 17 yielded a sensitivity of .79 and specificity of .63. Perhaps this could be attributed to the characteristics of the participants as noted by Deb et al. (2009); they screened for autism in children with intellectual disabilities only, while the Brereton et al. (2002) examined the validity of the tool among individuals with and without intellectual disabilities. Neither study reported a PPV or NPV. There was no validation against an accepted gold standard tool; rather, the participants received a clinical diagnosis of autism based on the ICD-10-DCR (International Classification of Diseases 10th Revision, Research Diagnostic Criteria) in the Deb et al. (2009) study.

Appraising the quality of the reviewed study (Deb et al., 2009), the content validity, structural validity, cross-cultural validity, internal consistency, construct validity and reliability were all rated as undetermined. Criterion validity was rated as sufficient based on the evidence. As peer-reviewed studies do not always provide sufficient information, the authors of the DBC were contacted to confirm which of the validities were examined. Based on the evidence provided by the authors and excerpts from the manual, reliability, internal consistency, convergent validity,

criterion validity, discriminative validity, and concurrent validity were all rated as positive. Since the DBC-ASA is not an independent measure but an algorithm within the DBC, the relevant psychometric (discriminative validity) property of the DBC-ASA was assessed. Brereton et al. (2002) and Deb et al. (2009) both reported that the DBC-ASA had very good discriminative ability. However, there remains inadequate information on cross-cultural validity, placing a limitation on its use in an African context. The overall rating for the tool based on the COSMIN checklist was medium.

4.4.2.6 Pervasive Developmental Disorder in Mentally Retarded Persons (PDD-MRS)

The PDD-MRS is a 12-item questionnaire designed for clinician screening for autism amongst those with intellectual disabilities. It has dichotomous items spread across the following domains: communication, social behaviour and stereotyped behaviour. It was designed to be used with children and adults ages 2 – 55 years old. The original Dutch version: the Autisme- en Verwante kontaktstoornissenschaal voor Zwakzinnigen (AVZ) was developed specifically for use with people with intellectual disabilities (Kraijer, 1990) with a revision in 1994 (Kraijer, 1994). The instrument is based upon the DSM-III-R criteria for pervasive developmental disorders and has been widely used in the Netherlands and Belgium.

Kraijer & de Bildt (2005) described and discussed the construction of the scale and its validation. The psychometric properties were tested on a sample of 1,230 participants with varying levels of intellectual disabilities. The resulting sensitivity at a cut-off score range of 10 – 19 was .92, while specificity was .92, but neither the PPV nor NPV was reported. Internal consistency for participants with functional speech was reported as α = .86 and for those without speech α = .81. Cortés et al. (2018) developed and validated the Escala de Valoración del Trastorno del Espectro Autista en Discapacidad Intelectual (EVTEA-DI), the Spanish version of the PDD-MRS. Reported results were r = .78 for convergent validity between the EVTEA-DI and the CARS, internal consistency measured by the Kuder-Richardson-20 (KR-20) was .71. At a cut-off score of 30, sensitivity was .71, specificity of .90, PPV = .73 and NPV = .90. To assess the discriminative validity of the EVTEA-DI, Cortés et al. (2018) utilised the Youden Index (YI). At a cut-off score of 8, sensitivity = .84 and specificity = .83.

For the PDD-MRS, content validity was rated as moderate based on the evidence from reviewed studies. Structural validity, internal consistency, criterion validity, and construct validity were all

rated as positive as there was sufficient methodological evidence found to support the rating. There was moderate evidence for cross-cultural validity since individuals with varying disabilities from different populations were participants. Studies were completed with Dutch and Spanish speaking participants. Reliability was rated as insufficient based on the COSMIN rating of lowest score counts. Authors were contacted for further evidence without success. The overall COSMIN rating for this tool was medium.

4.4.2.7 Child Behavior Checklist (CBCL)

The CBCL (Achenbach & Rescorla, 2001) is now a component of the Achenbach System of Empirically Based Assessment (ASEBA). The CBCL is a caregiver report questionnaire on which children and teenagers (2-18 yrs) are rated for various behavioural and emotional difficulties. Associated with disorders from the DSM-5, it measures difficulties on a scale made up of eight categories – rule-breaking behaviour, anxious/depressed, social problems, somatic complaints, thought problems, attention problems, withdrawn/depressed and aggressive behaviour. The form consists of 118 items that take between 30 minutes to an hour to complete. The CBCL has been translated into 60 different languages. Previous versions of the checklist were not designed to screen for autism in young children older than 4 years, and 6 years in the current revision (Mazefsky et al., 2011).

However, Ooi et al. (2011) aimed to derive and test an autism scale that could significantly differentiate children and adolescents with and without autism using the CBCL. The study participants were between 4 and 18 years old. The researchers considered whether eight scale factors could significantly differentiate individuals with and without autism, and they reported a sensitivity range of 48 – 78% and a specificity range of 59 – 87%. Following this, Ooi et al. (2011) derived and tested an autism scale comprised of items taken from the CBCL that significantly differentiated autistic children from other groups. Results showed that nine specific items were predictive of autism with sensitivity ranging from .68 – .78 and specificity range of .73 – .92. The PPV and NPV were not reported. The CBCL scores falling below the 93rd percentile is considered normal, scores between the 93rd to 97th percentile are borderline clinical, while scores above the 97th percentile are in the clinical range. Results of Ooi et al. (2011) are consistent with findings from previous studies (Mazefsky et al., 2011). Both Ooi et al. (2011) and Mazefsky et al. (2011) reported that the CBCL scales with more effective discriminative

abilities between the typical and autistic school-aged children were the 'Thought Problems, Social Problems and Withdrawn/Depressed' categories.

Regarding the quality appraisal from the reviewed study Ooi et al. (2011), the content validity for the CBCL was rated as indeterminate while structural validity was rated as positive, given the quality of the evidence reviewed. Criterion validity, construct validity and internal consistency were all rated as undetermined as there was not sufficient evidence. There was moderate evidence for reliability, with sufficient evidence to rate cross-cultural validity as positive. The scale which was originally developed in English was used with participants in three different languages (English, Malay and Tamil) and five different groups (Ooi et al., 2011). The authors were contacted for more evidence or access to relevant portions of the manual. Based on the author's response, content validity, reliability, criterion validity, construct validity, internal consistency and discriminative validity were all rated as sufficient. The overall rating for the CBCL was, medium, based on the level of evidence using the COSMIN checklist. Although work has gone into translating the tool into different languages and deriving a potential autism specific screening subscale from the CBCL, some training is required. The level of training depends on how the data are to be used. For LMICs such as Nigeria, Ghana, Kenya and other African countries, these requirements are potential barriers.

4.4.2.8 Mobile Autism Risk Assessment (MARA)

Duda et al. (2016) described the MARA, a new 7-item parent or caregiver questionnaire designed to screen for individuals at risk of autism. The MARA was developed based on the analysis of a pool of ADI-R score sheets of individuals with and without autism. An alternating decision tree algorithm was used to generate the questions and responses. The tool is administered and scored electronically, and the reported sensitivity was .90 and specificity was .80. Given that the data used for testing the measure were taken from the ADI-R, it should follow that the discriminatory ability and construct validity should be good. The MARA was validated against the Autism Diagnostic Observation Schedule (ADOS) and the PPV was .67, and NPV was .95. Duda et al. (2016) reported no specific cut-off scores; however, they referenced Wall, Dally, Luyster, Jung, & DeLuca (2012) where they used a categorical variable with two options – autistic or not autistic. Although the MARA looks promising, more large-scale reliability and validity studies with participants of differing developmental abilities are needed.

Based on the reviewed study (Duda et al., 2016), there was adequate evidence to rate structural validity as positive. Internal consistency, reliability, criterion validity, and construct validity were rated negative due to insufficient evidence. Content validity was rated as insufficient as the involvement of experts and users was unclear. Evidence for cross-cultural validity was insufficient and was rated as very low. The authors were contacted for more information and possible access to the manual if available. Based on feedback from one of the authors, content validity was revised to a positive rating. However, other studies provided were not on the MARA but on detecting ASD through Machine Learning. Participants in those studies were children younger than 5-years of age, thus not meeting the inclusion criteria for this review. Based on the COSMIN standard, the overall rating for the measure was low. Also, using this tool in Africa could be challenging, given that not everyone has internet access or personal computers.

4.4.2.9 EDUTEA: A DSM-5 Teacher Screening Questionnaire for Autism and Social Communication Disorders (EDUTEA)

The EDUTEA was developed in Spain as a brief autism screening tool for use by teachers and school professionals who had limited time (Morales-Hidalgo et al., 2017). The EDUTEA is an 11-item questionnaire based upon DSM-5 diagnostic criteria and was designed to enable teachers to gain information about the social interactions, behaviours and communication skills of children. The tool was validated against the ADOS-2, ADI-R and compared to the CBCL, Childhood Autism Spectrum Test (CAST) and Schedule for Affective Disorders and Schizophrenia (K-SADS-PL). Scoring of items is on a 4-point Likert scale, resulting in a minimum score of 0 to a maximum score of 33.

In evaluating the discriminatory ability and psychometric properties of the tool, Morales-Hidalgo et al. (2017) recommended a cut-off score of 10. At the recommended cut-off, the EDUTEA successfully discriminated between those with autism and related disorders and those with ADHD with an associated sensitivity of .83 and specificity of .73. For differentiating individuals at risk of autism or social pragmatic communication disorder (SCD), the authors reported good discriminatory abilities at the cut-off score of 10, with sensitivity = .87 and specificity = .91NPV of .99 and a PPV of .87. The two-factor internal consistency for the measure was α = .95 for social communication impairments and α = .93 for restricted behaviour patterns. Overall internal consistency was α = .97. No other studies using the instrument were found from the literature search.

Content validity was rated as positive as teachers were involved in the development of the instrument. The structural validity, internal consistency, criterion validity, reliability and construct validity were all positive with moderate evidence. However, cross-cultural validity was judged as having insufficient evidence. The overall rating based on COSMIN standards was medium.

4.4.2.10 Autism Diagnostic Inventory-Telephone Screening in Spanish (ADI-TSS)

Vranic et al. (2002) developed the ADI-TSS as a semi-structured interview administered over the telephone. ADI-TSS was modelled upon the Autism Diagnostic Interview-Revised (ADI-R) with forty-seven questions in three areas. The final version used in the study was administered to 59 participants and had a sensitivity of 1, and a specificity of .66 with no PPV or NPV reported. Although this tool was developed over fifteen years ago, no other studies validating its use and properties were found. Interrater reliability for the subscales were as follows: social reciprocity $\alpha = .94$, verbal communication $\alpha = .93$, non-verbal communication $\alpha = .94$, and repetitive behaviour $\alpha = .94$.

Content validity for the subscales was rated positive, while the overall content validity was rated low due to insufficient evidence for end-user input in the development of the tool. Structural validity and internal consistency were rated insufficient. Cross-cultural validity was rated insufficient as the translation methodology was unclear. Although inter-rater reliability for the subscales was shown, there was insufficient evidence for the reliability of the total tool; thus, this was rated insufficient. Based on the COSMIN checklist, the tool was rated as low overall. The feasibility of using the ADI-TSS in Africa, where there are high costs associated with mobile telephone use would be a challenge.

4.4.2.11 Diagnostic Behavioral Assessment for Autism Spectrum Disorder-Revised (DiBAS-R)

The DiBAS was developed by Sappok and colleagues (2014b) to help with screening autism amongst adults with intellectual disabilities. It was designed to be administered by caregivers or individuals knowledgeable about the person, but who also lacked specific knowledge about autism. The 20-item questionnaire was derived from the ICD-10 and DSM-5 criteria for autism. To improve its diagnostic validity further, a single item was deleted following a pilot study and item-revision of the DiBAS (Sappok et al., 2014a). The resulting 19-item screening tool can be

completed in 5-minutes by a caregiver, family member, staff or any person who is familiar with the individual.

Heinrich et al. (2018) assessed the diagnostic validity of the DiBAS-R in 381 adolescents and adults with intellectual disabilities, some of who had autism. Study participants ages ranged between 16 – 75 years. Based on the recommended cut-off score of 29, the reported results were sensitivity = .82, specificity = .67, the PPV = .44 and the NPV = .92. The participant's diagnosis was confirmed using the ADOS and ADI-R.

Based on the reviewed study (Heinrich et al. (2018), content validity was rated as undetermined. Expert clinicians participated in the development, but the item reduction process was unclear. Assessment of comprehensibility and comprehensiveness was also unclear. Evidence for cross-cultural validity, structural validity and internal consistency were also insufficient. Reliability was rated as insufficient, while criterion validity and construct validity had sufficient evidence to rate them positive. The authors were contacted for access to the manual or further evidence on the tool's development. Since the manual is in German, the authors provided Sappok et al. (2014a) in which the relevant information was reported. Following this, the content validity, structural validity, internal consistency, reliability, convergent and discriminative validity were all rated positive. However, evidence for cross-cultural validity remained insufficient. DiBAS-R was rated as medium based on the additional evidence using the COSMIN checklist. DiBAS-R is currently available only in German, thereby limiting the feasibility of using it in Africa.

4.4.2.12 Adapted Autism Behaviour Checklist (AABC)

The AABC which is based on the Autism Behavior Checklist (Krug et al., 1980) is a 57-item measure developed in Turkey by Özdemir & Diken (2018). Modifications were made to the original form to include the ICD-10 and DSM-5 criteria for autism. The measure was designed to be completed by a parent, primary caregiver, or a teacher familiar with the individual and then scored and interpreted by a trained professional.

Özdemir & Diken (2018) assessed the diagnostic validity of the AABC in 1,133 children and adolescents with autism and intellectual disabilities. Study participants ages ranged between 3 – 15 years. Reported results were r = .73 between the AABC and the Gilliam Autism Rating Scale-2 Turkish Version (GARS-2 TV), internal consistency measured by the Kuder-Richardson-21

(KR-21) was .89, test-retest reliability was r = .82 and correlation between the two-factors (social limitations and problematic/repetitive behaviours) was r = .46. At a cut-off score of 13, the measure discriminated between the ASD and ID groups reliably with a sensitivity of .87 and specificity of .82.

Based on the COSMIN checklist, content validity, structural validity, internal consistency, reliability, criterion validity and construct validity were all rated as positive. Cross-cultural validity was rated as insufficient based on the evidence. The tool has only been used in Turkey. Since this measure is only available in Turkish, the feasibility of using it in Africa is limited as substantial resources would be required for translation. The overall rating for the measure was medium.

4.5 Intellectual Disability

The ninteen studies identified focused upon people with intellectual disabilities and included a total of 3,129 participants with age ranging from 3 to 74 years. Like autism, studies with participants younger than 11-years or older than 26-years old were included when some or the majority of their participants were within the specified age range of the inclusion criteria. The number of studies by country was as follows: UK (n = 7), USA (n = 4), Norway (n = 5), and one each from Australia, Netherlands, and Belgium (Table 9). Three of the studies (McKenzie et al., 2012b; Trivedi, 1977; Ford et al., 2008) involved adolescents only while fifteen studies involved a combination of children, adolescents, and adults. The screening tools used in the studies were the Slosson Intelligence Test (SIT), Learning Disability Screening Questionnaire (LDSQ), Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q), Screener for Intelligence and Learning Disabilities (SCIL), Hayes Ability Screening Index (HASI), and the Quick Test (QT). Validation of these screening tools was against full-length tests considered as the gold standard, such as the), different editions or versions of the Weschler scales and the Adaptive Behavior Assessment System-Second Edition (ABAS-II). In some instances, screening tools were compared to other full-length scales, such as the Stanford-Binet Intelligence Scale, or to similar short measures. For example, the HASI was compared with the KBIT and the SIT with the Peabody Picture Vocabulary Test (PPVT).

The quality ratings for the included studies are found in Tables 6 and 7. Five of the included studies made use of a single group of participants, while eight used a between-group design, and six a within-group design. Each screening tool is considered in turn below.

4.5.1 Intellectual Disabilities Screening Tools

4.5.1.1 Slosson Intelligence Test (SIT) and Slosson Intelligence Test-Revised (SIT-R)

The original SIT was developed by Richard Slosson in 1963 and used as part of an assessment to determine whether an individual has an intellectual disability, measured as IQ. At the time of this review, no studies utilising the third and fourth versions of the SIT were found. Rotatori and Epstein (1978) assessed the ability of special education teachers without previous psychological testing experience, to reliably administer the SIT. Reported test-retest reliability results (r = .94) appeared excellent, indicating that the test was reliable over time when administered by special education teachers. To examine the concurrent validity of the revised SIT, Kunen et al. (1996) compared the SIT-R to the Stanford-Binet Intelligence Scale, Fourth Edition. The correlation was high (r = .92), but the consistency of the IQ classification between the two instruments for those who had intellectual disabilities was poor. In comparison to the Stanford-Binet Intelligence Scale, the SIT-R had insufficient evidence of construct validity due to discrepancies in match rates between the SIT and the Stanford-Binet. For instance, for the entire study sample with Iqs ranging from 36 to 110 (Kunen et al., 1996), there was a 50% match rate between the Stanford-Binet and the SIT for all the classifications, mild, moderate, average and low average Iqs. Nevertheless, for the mild to moderate categories out of the 38 participants categorised as mild on the Stanford-Binet, SIT categorised them as 1- low, 2- slow, 9- mild, and 26- moderate. Trivedi (1977) meanwhile examined the comparability of the SIT against the Wechsler Intelligence Scale for Children (WISC) in adolescents. He found significant correlations between the WISC and SIT when compared on mental age (r = .87) and IQ (r = .86). Trivedi (1977) concluded that the SIT reliably approximated the WISC as a screening tool.

Blackwell and Madere (2005) commented that the SIT-R demonstrated and fulfilled its stated purpose of "being a valid, reliable, individual screening test of general verbal cognitive ability" (p. 184) but have also suggested problems with the reliability and validity of the SIT-R. Reviews by other authors have also raised concerns about the reliability and validity of the SIT-R

(Campbell & Ashmore, 1995). Potential challenges regarding the use of the SIT-R with those from multicultural backgrounds, or where English is a second language were reported by Blackwell & Madere (2005). Other limitations of the SIT-R are its inability to measure functioning levels of other intellectual areas such as perceptual-motor functioning. There is also the difficulty of comparing SIT scores with those of other IQ tests for persons older than 16-years of age due to the unclear and insufficient methodological information given by the developers (Campbell & Ashmore, 1995). Although the SIT has the above limitations, one advantage is that persons with limited psychometric training and knowledge can administer it.

Based on the COSMIN checklist, the SIT (or SIT-R) was rated as low overall. There was sufficient evidence for reliability from the studies reviewed for it to be rated as moderate. Content validity and structural validity were rated undetermined. Both criterion validity and construct validity were rated as inconsistent. Internal consistency and cross-cultural validity were rated as negative, based on the poor amount of evidence.

4.5.1.2 Quick Test (QT)

The QT is an intelligence test measuring verbal information processing and receptive vocabulary (Ammons and Ammons, 1962). It comprises three parallel forms with 50 items, each of which can be administered to children and adults. Verbal intelligence is measured by the ability to match words of increasing complexity to pictures. Sawyer & Whitten (1972) investigated the concurrent validity of the individual and combined scores of QT against the WISC sub-tests. Moderate correlations (r = .33 – .52) were reported for the picture arrangement, coding, performance scale score and the full-scale score. For the verbal scale, the correlation between both the QT and the WISC was between r = .31 and .34 for both the individual and combined forms of the QT. One challenge with the QT is that it predominantly measures verbal skills. This limitation may have impacted the Sawyer & Whitten (1972) study, as most of the participants had limited verbal ability. Moreover, the pictures used are rather old-fashioned and may not transfer well to the African context.

Based on the COSMIN checklist, the overall evidence for the QT was very low. Structural validity, internal consistency and reliability were rated low based on insufficient amount of evidence both from the study and manual. There was sufficient evidence to rate the construct

validity, content validity and criterion validity as positive, while cross-cultural validity was undetermined. The overall rating for the QT was very low.

4.5.1.3 Hayes Ability Screening Index (HASI)

The HASI is a brief screening tool for intellectual abilities comprised of four subtests covering background information, puzzle, clock drawing and backward spelling (Hayes, 2000). The HASI has been used predominantly in criminal justice settings to identify vulnerable persons with intellectual disabilities. HASI is designed for use with people aged 13 to adulthood. For those aged 13 – 18 years, the cut-off score is 90, while for those older than 18-years, it is 85. Some training is required before its use.

Hayes (2002) reported on the construct validity of the HASI and the correlation with the Kaufman Brief Intelligence Test (KBIT), Wechsler Abbreviated Scale of Intelligence (WASI) and WISC-III. The total population sample correlation between the HASI and KBIT was reported as high (r = .62). The reported sensitivity for the study was .82, and the specificity was .72. Hayes (2002) suggested that the youth cut-off be maintained at 90. A different study (Ford et al., 2008) which had all adolescent (10 – 19-year-olds) participants, found a correlation of r = .55 between the HASI and the FSIQ of the WISC-IV or the Wechsler Adult Intelligence Scale (WAIS-III) and r = .38 with the Vineland Adaptive Behavior Scale (VABS). At the recommended cut-off score of 90 for those below 18 years of age and 85 for those over 18 years old, the authors reported a poor agreement (k = .25) between the HASI and the FSIQ from the Weschler scales when categorising as ID. Sensitivity at these cut-off scores was .66 and specificity of .51. Lowering the cut-off score to 80.2 yielded better agreement (k = .54) a sensitivity of .80 and specificity of .65. Søndenaa et al. (2007) translated the HASI to Norwegian and validated the construct and criteria of the screening tool against the Norwegian version of the WAIS-III. The study participants were between 17 and 60 years old. The authors found a high correlation between both instruments (r = .81) with an internal consistency of α = .76. Søndenaa et al. (2007) also reported that scores on all HASI subtests, WAIS-III FSIQ and the verbal and performance subscales were significantly correlated with r above .61. At the recommended cut-off score of 85 for indicating ID, the sensitivity was 1 and specificity .57. However, Søndenaa et al. (2007) adjusted the cut-off score to 81 for their sample to reduce the over-inclusion of false positives. The alternative cut-off of 81 yielded a sensitivity of .95 and specificity of .72.

In the Søndenaa et al. (2008) prevalence study, the HASI was validated against the WASI as a screening tool. The HASI was found to be somewhat overly inclusive with a specificity of 72.4% and sensitivity of 93.3%. Correlations between the WASI full-scale and HASI were significant with r = .72, verbal tests r = .63 and performance tests r = .74. In Søndenaa et al. (2011), the criterion validity of the HASI was examined against the WASI with a psychiatric population. The study reported the over categorisation by the HASI with a sensitivity of 1 and specificity of .35 at the recommended cut-off score as previously mentioned. However, the authors argued that the HASI is designed to be overly inclusive, since it is better to identify everyone who may need full assessments, rather than miss some people. Also, Søndenaa et al. (2011) reported moderate correlations between the subtests of the WASI and HASI (r = .67). However, when the "background information" subtest was eliminated, correlation increased to r = .71 and internal consistency of α = .67.

To et al. (2015) examined the discriminative and convergent validities of the Dutch version of the HASI against the WASI-III in persons with substance abuse problems. Convergent validity between the HASI and WAIS-III FSIQ scores, were significantly correlated (r = .69). There was also a correlation between the HASI subtests and the WAIS-III as follows: background information r = .58, spelling r = .50, puzzle r = .46, clock drawing r - .45, verbal subscale r = .70, and the performance subscale was r = .63. Discriminant validity was reported as significant from the receiver operating characteristics (ROC), with an area under the curve (AUC) of .95 yielding a sensitivity of .91 and specificity of .80 at the cut-off score of 85. In Braatveit et al. (2018), the convergent and discriminative validities of the Norwegian version of the HASI were examined in a population of persons with a substance abuse history. At the cut-off of 85, sensitivity was reported as 1 and specificity of .65. Braatveit et al. (2018) also reported that lowering the cut-off score to 80.7 yielded increased specificity of .81 without affecting the sensitivity. Similar to Søndenaa et al. (2011), Braatveit et al. (2018) also mentioned that the over-categorisation by the HASI was intended to be a means of detecting other persons with/without intellectual disabilities but who may benefit from further evaluation. Regarding convergent validity, Braatveit et al. (2018) correlated the HASI against the full-scale WAIS-IV with a significant correlation (r = .70).

Based on the reviewed studies, and the COSMIN checklist, the overall rating for the HASI was low. Reliability was rated as negative due to insufficient evidence. Structural validity had inadequate evidence and was rated as undetermined. Content validity was rated as low due to insufficient evidence. Criterion validity and construct validity were rated positive with excellent evidence. There was moderate evidence for a positive rating on the cross-cultural validity based on the use of the Norwegian and Dutch versions, as well as the original Australian version. To ensure that all relevant properties of the tool were properly rated, the manual was consulted. Based on the manual, additional ratings employing the COSMIN were made. Content validity remained low as there was no evidence on expert clinicians or end users involvement in the development. Evidence for internal consistency was not in the manual thus a rating of insufficient was given. Reliability was rated as insufficient, while criterion validity and construct validity had sufficient evidence to retain their positive rating. The overall rating of the HASI was revised to medium following the combined evidence from the studies and the manual. Although the HASI has been adapted for use in two further languages and environments outside of the original development area, most of the studies used the tool in the Criminal Justice System. Studies that employed the tool with adolescents outside of the CJS would have been more useful for forming a decision on adapting it for use in Africa and countries like Nigeria.

4.5.1.4 Learning Disability Screening Questionnaire (LDSQ)

McKenzie and Paxton (2006a) developed this 7-item screener for the identification of adults with intellectual disabilities to assist in deciding eligibility for community services. The LDSQ has also been used in criminal justice and forensic settings. Areas assessed include literacy, living situation and employment. The LDSQ has been reported to have both criterion and convergent validity when compared to the WAIS-III (McKenzie and Paxton, 2006b). McKenzie et al. (2012b) examined the convergent and discriminative validities of the LDSQ in forensic settings. Convergent validity between the FSIQ and the LDSQ was reported as highly significant with a correlation coefficient of $r = .71$. The authors also reported good discriminative ability of the LDSQ with a sensitivity of 82.3% and specificity of 87.5% based on the receiver operating characteristics analysis (AUC = .898). PPV and NPV were reported as 92.9% and 73.7% respectively. McKenzie et al. (2015) validated the LDSQ's criteria against a standardised tool, the WAIS-IV FSIQ and reported a good correlation between them with a sensitivity of .92 and specificity of .92 (AUC = .945). Convergent validity was reportedly significant for the WAIS-IV

FSIQ and LDSQ total performance with a coefficient of $r = .71$. Significant correlations were also reported for the subtests – verbal comprehension ($r = .54$), perceptual reasoning ($r = .69$), working memory ($r = .58$), and processing speed ($r = .58$). Although these studies by McKenzie et al. (2012a; 2014) reported excellent psychometric properties for the LDSQ, the independent study by Stirk, Field & Black (2018) reported a sensitivity of .67 and specificity of .71 at the threshold given by McKenzie et al. (2014), showing that the LDSQ may require more investigation to align the properties with recommended standards (Glascoe, 2005).

Based on the evidence from the studies reviewed, criterion validity and construct validity were rated as moderate. Content validity was rated as insufficient since there was not enough evidence of user participation in the development of the tool. Structural validity, internal consistency, cross-cultural validity, reliability, were all rated low due to insufficient evidence. The manual was obtained to confirm which of the tool's properties were examined during development. From the manual, there was moderate evidence for content validity, discriminative validity, and convergent validity. Interrater reliability was assessed while there was no evidence for internal consistency. Combining the evidence from the studies and the manual, the overall quality of the LDSQ was rated as medium using the COSMIN checklist. Like the HASI, this measure has been used primarily with adults in the CJS and forensic services. However, unlike the HASI, evidence to support the cross-cultural application was not apparent, and so the feasibility of its use with African adolescents is limited.

4.5.1.5 The Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q)

The CAIDS-Q, modelled after the LDSQ, was developed by McKenzie and Paxton in 2012 as a short 7-item screening questionnaire for detecting intellectual disabilities in children and adolescents in mental health and forensic services. It is designed for use with individuals aged 8 – 18 years. According to McKenzie & Paxton, the CAIDS-Q can discriminate between those with and without intellectual disabilities with 97% accuracy. Four studies, McKenzie et al. (2014; 2012b, 2012c; 2019) evaluated and validated the psychometric properties of the tool and reported values within recommended standards. McKenzie et al. (2014) assessed the discriminatory ability of the CAIDS-Q against a short form of the WISC-IV with a sample of children aged 10 to 11 years with and without intellectual disabilities (who had been fully assessed for this on either the WISC IV or the WAIS III). Overall, the WISC-IV short form itself

led to the correct classification of 91% of the participants. When broken down, the classification of those with intellectual disabilities was 92% correct while those without intellectual disabilities was 91% correct, using the WISC-IV short form. AUC for the WISC-IV was .98 which gave a PPV of .87 and NPV of .95. The CAIDS-Q led to the correct classification of 89% of children with intellectual disabilities, and 88% of those without; the PPV was .92, and the NPV was .85 based on an AUC = .94. Overall, the CAIDS-Q correctly classified 88% of the participants. McKenzie et al. (2012c) evaluated the face, construct, criterion, convergent and discriminative validity of the CAIDS-Q with comparisons made to either the WISC-IV FSIQ or the WAIS-III FSIQ depending on the participant's age. Results obtained from the study showed high internal consistency ($\alpha = .88$), significant correlations between the CAIDS-Q and the WISC-IV FSIQ ($r = .78$), and significant correlations between the CAIDS-Q and the WAIS-III FSIQ ($r = .79$). At a cut-off of 62 for the children (8 – 11 years), the measure had a sensitivity of .97 and specificity of .86, and at a cut-off of 64 for the adolescents (12 – 18 years) the sensitivity was .96 and specificity .85. McKenzie et al. (2012c) reported that there was no significant difference between age and the CAIDS-Q score for the total population ($r = .02$). McKenzie et al. (2012b) evaluated the convergent and discriminative validity of the CAIDS-Q against the WISC-IV in a forensic setting. Reported outcomes were significant correlations between the CAIDS-Q and the FSIQ ($r = .76$), with correlations between the CAIDS-Q score and the subtests as follows: verbal comprehension ($r = .54$), perceptual reasoning ($r = .65$), working memory ($r = .52$), and processing speed ($r = .74$). Other results include a PPV of 1, NPV of 1 and good internal consistency ($\alpha = .72$). McKenzie et al. (2019) examined the convergent validity, test-retest reliability, interrater reliability, sensitivity, specificity, PPV and NPV in a paediatric neurodevelopmental setting based on previously determined cut-off scores. Convergent validity of the CAIDS-Q was examined against the WISC-IV and Adaptive Behaviour Assessment System, Second/Third Edition (ABAS II/III). Reported correlations between the CAIDS-Q and the FSIQ ranged between $r = .62$ - $.79$, with correlations between the CAIDS-Q and the ABAS GAC ranging between $r = .48$ - $.60$. Other results include a PPV of 1, NPV of .78, sensitivity = 1 and specificity = .88 for the total sample. A two-week time frame yielded a test-retest correlation of $r_{32} = .90$ while interrater reliability *k* was between .26 and 1 for the four items (time, read, write and laces) tested.

From the studies reviewed, content validity was undetermined as user participation in the development was unclear. Assessment of comprehensibility was also unclear. Structural validity was rated as negative due to insufficient evidence. Evidence for reliability was moderate from the studies. Cross-cultural validity was rated as moderate since the measure was used with two different age groups: children and adolescents. Criterion validity was also rated as moderate. However, incorporating information from the manual, content validity, structural validity, internal consistency, reliability, criterion validity and construct validity were all rated as positive with moderate evidence. Based on the COSMIN checklist, the overall rating of the tool was medium.

4.5.1.6 Screener for Intelligence and Learning Disabilities (SCIL)

The SCIL is a tool for identifying persons with a level of general intellectual functioning that falls within and below the "borderline" range (Nijman et al., 2018). The SCIL comprises elements of social adaptive skills, language comprehension, education, arithmetic, reading and writing abilities. Geijsen et al. (2016) examined the predictive validity of the SCIL for identifying intellectual disabilities amongst adolescents and adults in a criminal justice setting (police detention). Reported results from the study showed that the SCIL total score correlated moderately with the WAIS-III short form (r = .56) with a sensitivity of .72 and specificity of .70. The PPV and NPV were not reported. Additionally, reliability was reported as α = .64 and α = .84 in a previous study (Kaal, Nijman, & Moonen, 2015). Nijman et al. (2018) conducted two further studies, split into adults and adolescents, and investigated the predictive validity of the SCIL. Participants in both groups included persons with and without intellectual disabilities. At the suggested cut-off of 19, sensitivity was .82 and specificity was .89 for adults, and for adolescents the suggested cut-offs varied according to age. For those aged between 16 and 17-years old, a cut-off score of 18 resulted in a sensitivity that was .80 and specificity that was .84; for those aged 14 to 15-years old, a cut-off score of 16 resulted in a sensitivity of .85 and specificity of .82. The AUC for adolescents as a total group was .91, and .93 for adults. The SCIL had test-retest reliability of r = .92. Nijman et al. (2018) analysed the internal consistency using the split-half method which yielded high correlations; α = .84 in the first half and α = .82 in the second half.

Based on the COSMIN checklist, the overall rating for SCIL was moderate. Content validity was rated positive with moderate evidence based on the involvement of experts and end-user. Structural validity, criterion validity and internal consistency were all rated as highly positive with enough evidence. Reliability had moderate evidence with consistent findings in both studies. Cross-cultural validity was positive as participants were recruited from different cities, police stations and refugee sites. The SCIL showed promising results, but more studies to validate the tool are required.

4.6 Discussion

Identification and selection of a user friendly, accessible, time-efficient, and useful screening tool for use in Africa requires careful thought and consideration. The focus of this review was on identifying potentially useful tools for screening African adolescents and younger adults with autism and/or intellectual disabilities. This age range was the focus as many who have intellectual disabilities and/or autism are noticed as they become more independent and when they begin to interact more often with others outside of their immediate family, for example in secondary school settings and the wider community. In interacting with these environments, disabilities and challenges become more obvious. Such adolescents may not have received a diagnosis earlier in their life because of a lack of awareness, insufficient or inadequate resources, limited numbers of professionals and the families sometimes not seeking immediate help for those individuals with autism or intellectual disabilities till later in life. To begin to address this gap, appropriate and suitable screening tools need to be designed or identified for use in Africa. Towards this, the aims of this review were to (1) describe and critically appraise a range of short screening tools for the detection of intellectual disabilities and autism, (2) consider the psychometric properties of these tools, and (3) consider the appropriateness of using these tools across a range of cultures. A discussion of the review findings is presented below.

4.6.1 Description and appraisal of short screening tools

A total of 12 tools screening tools for autism were identified through this review. The tools are the ADI-TSS, EDUTEA, PDD-MRS, DiBAS-R, AQ-10, ASSQ-REV, SCQ, CARS, CBCL, DBC-ASA, AABC and the MARA. Apart from the AQ-10 adolescent version, all the other tools were designed to be used across a wide age range. The CARS and CBCL were not originally

developed as screening tools, but the studies (Mesibov et al., 1989; Ooi et al., 2011) reviewed utilised them as such with the intent of developing subscales for autism screening. Moreover, the CBCL has over 100 items and takes between 30 minutes to an hour to complete, which goes against the timing for brief tools. Both the CARS and CBCL require some training and specific qualifications before use. Given the socioeconomic climate of African countries, the resources required to gain specific administrative qualifications for these tools may not be readily and widely available. As such, there will be challenges associated with routine use within Africa.

Both the PDD-MRS and DiBAS-R can be used across a wide age range from 2 – 80-year-olds and administered in 5 to 20 minutes. The wide age range allows for their use with adolescents while the short administration time qualifies them as short and time-efficient tools. The PDD-MRS and DiBAS-R were designed for use with persons who are known to have intellectual disabilities. Limiting the measures to those with known intellectual disabilities presumes those individuals have been diagnosed; this is not entirely the situation in Africa. Considering this design limitation, the feasibility of their use in Africa will be challenging.

The MARA, ADI-TSS and SCQ were modelled after the ADI-R. While the SCQ and MARA take between 5 – 10 minutes to administer, the ADI-TSS takes between 20 – 40 minutes. The lengthy administration time of the ADI-TSS may be because of the telephone administration. As an over-the-telephone screening tool, the usefulness of ADI-TSS in Africa, where not everyone may have access to a telephone, poses immediate limitations. Similar constraints are associated with the MARA, which is a computer-based parent or carer administered screening tool. The number of persons with immediate access to either a smart device, personal computer or constant electricity is likely to be low in the African continent or individual African countries. This lack of immediate access to smart devices poses a limitation to the usefulness of the MARA in Africa. Meanwhile, for the SCQ, seven (32%) out of the 22 studies reviewed employed this tool and this observation is consistent with findings from previous studies that concluded the SCQ was used more widely in research (Bozalek, 2013). The SCQ comprises two forms: lifetime and current. The SCQ current form is used to assess an individual's behaviour during the past 3 months while the lifetime form assesses the developmental history. One advantage of the SCQ is the availability of the lifetime form, which enables information gathering for adolescents who have

never been screened. This feature, amongst others, makes the SCQ a viable option for use with African adolescents.

The DBC-ASA, which is a subset of the Developmental Behavior Checklist, is limited to those under age 18-years, which is about the midpoint of the adolescent age range (11 – 26 years). The upper age limit of the screening tool poses a current challenge for routine use of the tool. On this basis, adopting the tool for use in Africa does not seem practicable without further standardisation work inclusive of a wider age range.

Both the ASSQ and AQ were developed for persons with HFA. While one of the reasons tools are developed is to bridge a gap or meet a need, in the African setting where screening is still in its infancy, using such disability-specific tools will not yield optimal results. EDUTEA, an 11-item questionnaire, was developed for use by teachers and school professionals. The study by Morales-Hidalgo et al. (2017) did not provide any information on administration time. Similarly, there was no information on administration time provided for the AABC, a 57-item questionnaire developed to be completed by parents, teachers, primary caregivers or persons familiar with the individual (Özdemir & Diken, 2018). Estimating the administration time based on the 11 or 57 questions introduces subjectivity when compared to the SCQ, which has 40 questions and takes 10 minutes. The EDUTEA and AABC are emerging tools and having more comprehensive information would have aided in forming an opinion about their usefulness in Africa.

For intellectual disabilities, a total of 6 tools were identified: the HASI, LDSQ, CAIDS-Q, SIT, SCIL and QT. Two of these tools (SIT and QT) focus solely on IQ scores to determine the presence of intellectual disabilities. Moreover, the QT is rather outdated and also measures mostly verbal skills, based on old-fashioned pictures which may not be culturally relevant to African settings. For individuals not verbally able, in Africa, the QT will not be very useful. The original SIT was considered outdated and not on a par with the Wechsler scales (Kunen et al., 1996) and was revised to address some of the concerns. However, the new SIT-R still focuses on verbal cognitive ability. In addition, as previously mentioned, other reviews of the SIT-R have mentioned problems associated with the reliability and validity of the tool. The LDSQ, meanwhile, is adult-specific, and studies that used the LDSQ had participants aged 18-years and above; 18-years is considered the legal adult age in most developed economies. The challenge

posed lies in the lower age limit of 18, implying that the LDSQ cannot be used with persons younger than 18-years old.

To close the gap, the CAIDS-Q was developed, by the authors of the LDSQ. CAIDS-Q is used for 8 – 18-year-olds. For screening adolescents, as defined by age 11 – 26 years, a more encompassing single measure is required. Two screening tools met this criterion, the HASI and the SCIL. HASI can be used with persons as young as 10-years, as there are two different cut-off scores: one for those below 18-years and another for those above 18-years. Given that HASI requires some training to use it and is also used largely in the CJS and forensic services, two areas that are underrepresented in the African context, these may impact on its usefulness in the African environment. The SCIL was developed and examined with adults (18 – 63 years) and adolescents (12 – 17 years). The SCIL also incorporates test items that assess social adaptive skills in line with the current diagnostic criteria for intellectual disabilities, as per the DSM-5, but is currently only available in the Dutch language.

A combined total of 18 screening tools were reviewed for autism and intellectual disabilities. The quality of the tool's design, studies employing them, and overall evidence provided were analysed using the COSMIN Risk of Bias checklist (Tables 2 through 7). Examples of the areas analysed were the concept elicitation, clearly describing the construct of interest, target population, and context of use. Based on the results of the review using the COSMIN checklist and additional information from manuals, the overall ratings for twelve tools (SCQ, CARS, PDD-MRS, EDUTEA, AABC, DiBAS-R, DBC-ASA, CBCL, LDSQ, CAIDS-Q, HASI and SCIL) were moderate. For four tools (AQ, MARA, SIT/SIT-R, and ADI-TSS) the rating was low, and the remaining 2 tools (ASSQ, and QT) were rated as very low.

4.6.2 Psychometric properties

For autism, clinical samples with a previous diagnosis participated in most of the studies, leading to a focus upon discriminative validity, differentiating those with and without autism. Using clinical data also meant that a comparison of the outcomes from the screening tools was not necessarily compared to those of an acceptable gold standard instrument. Regarding sensitivity, specificity, PPV and NPV, most of the studies reported values for specificity and sensitivity only. Psychometric properties from the studies reviewed were quite varied (Table 8). The

variations could be due to the heterogeneity of the participants across age, gender, severity, or the adjustment in cut-off scores. Other factors that can impact outcomes are study methodology and proxy informants (Ehlers & Gillberg, 1993). Deriving a cut-off score that is associated with precision is part of the development of instruments; however, in some studies, these adjustments resulted in marked variations. This variability was exemplified in studies that utilised the LDSQ. The studies by McKenzie et al. (2012a, 2015) reported sensitivities of .82 and .92, respectively, while Stirk et al. (2018) reported a sensitivity of .67.

Applying the guidelines from the COSMIN checklist, the quality of studies on measurement properties and the evidence for those properties were analysed (details in Tables 2 and 3). The properties included content validity (this includes relevance of the items in the tool, comprehensiveness, and comprehensibility), structural validity, internal consistency, reliability, criterion validity and construct validity. In rating the content validity, expert and end-user input are considered. COSMIN ratings are based on the 'lowest score' counts, as previously mentioned, and this formed the basis for the overall rating of the studies and outcomes. Eight tools (PDD-MRS, DiBAS-R, CBCL, DBC-ASA, MARA, AABC, SCQ & EDUTEA) had moderate evidence for content validity with the remaining four rated low. There was moderate evidence for structural validity for eight tools (SCQ, CARS, CBCL, DBC-ASA, MARA, EDUTEA, DiBAS-R & PDD-MRS). AABC had high evidence for structural validity while the remaining 3 had low or very low evidence. Only 82% (18) of the studies examined criterion validity, and these were studies that used the EDUTEA, PDD-MRS, ADI-TSS, DiBAS-R, AQ-10, ASSQ, SCQ, AABC and DBC-ASA. Out of these, the PDD-MRS, EDUTEA and SCQ were rated high while the DiBAS-R, AABC and ADI-TSS were rated moderate. Evidence from the remaining three tools was inadequate, and they received ratings of low. There was enough evidence to give a rating of moderate to the ADI-TSS, EDUTEA, SCQ, and DiBAS-R for construct validity while the PDD-MRS received a rating of high. Reliability was high in the EDUTEA and moderate for SCQ, CBCL, and CARS. Internal consistency was found to be high in the EDUTEA and moderate for the SCQ, PDD-MRS, AABC, CARS, DiBAS-R and AQ-10. Some ratings for the DBC-ASA and the CBCL were based on the manuals not on the studies. When all components of the psychometric properties are considered, the SCQ, CBCL, DBC-ASA, and PDD-MRS met most of the COSMIN criteria.

Turning to consider intellectual disabilities, out of the nineteen studies reviewed, fifteen of them which used the HASI, CAIDS-Q, LDSQ and SCIL incorporated the current DSM-5 criteria for intellectual disabilities by using both IQ and adaptive behaviour in the tools. The other four based on the SIT and QT focused on making comparisons with Full-Scale IQ as a basis for identifying participants with intellectual disabilities. Seventeen of the studies reviewed (89%) validated their results against the age-appropriate Wechsler scales, the most widely used assessment of general intellectual functioning, and often regarded as the gold standard. One study (Kunen et al., 1996) compared the SIT to the Stanford-Binet while Rotatori & Epstein (1978) focused on test-retest reliability.

All studies involving people with intellectual disabilities had evidence of explicit constructs for the development of the tools and, like the autism studies, these studies examined mainly the discriminative and predictive validities of the measures. Criterion validity was examined in all the studies with the HASI, and SCIL rated as high; while those with SIT, CAIDS-Q, LDSQ and QT were rated moderate. Evidence for construct validity was high for the HASI and CAIDS-Q and moderate for SIT, SCIL, LDSQ and QT. Internal consistency was high in the SCIL, moderate for CAIDS-Q and low for the HASI, LDSQ and QT, and very low for the SIT. The quality of evidence for content validity was moderate for the SCIL, CAIDS-Q, LDSQ and QT while very low for SIT and low for the HASI.

Regarding reliability, the HASI, QT and LDSQ were rated low while the SCIL, SIT and CAIDS-Q were moderate. Structural validity was rated high for the SCIL, moderate for CAIDS-Q and low to very low for the remaining four. Sensitivity, specificity, PPV and NPV values from the studies were also varied but generally within acceptable ranges (Glascoe, 2005). Sensitivity was between .67 and 1 while specificity was between .35 and .92. Based on this review, none of the intellectual disabilities screening tools identified through this review seemed to have been used in Africa. The SCIL and CAIDS-Q were found to have better overall psychometric properties and scored better on the COSMIN checklist (Tables 3 and 7). Not all studies incorporated adaptive behaviour scores alongside IQ and overall, in the future, there needs to be more of a shift from IQ testing as a measure of intellectual disabilities, to incorporating adaptive skills during screening and eventual diagnosis by using a tool that captures both.

4.6.3 Cultural adaptation

A key element for any of the tools selected for use within African nations is cross-cultural validity. Cross-cultural validity based on the COSMIN checklist includes the sample size, agreement between the original and translated versions, use with different populations, diagnoses and ethnicities. For example, a Spanish version compared to an English version, or Dutch participants compared to German participants or adults to adolescents.

All the screening tools identified through the review were used for both male and female participants. Comparisons were made between those with and without autism or intellectual disabilities Concerning the use of different versions of each tool, the AQ in English was used only in the UK, the English ASSQ in the USA and the Swedish version in Sweden. DiBAS-R which is in German was used in Germany, MARA in the USA, ADI-TSS Spanish version was developed and used in Argentina, PDD-MRS in the Netherlands where it originated as well as the Spanish version used in Spain, and the EDUTEA in Spain where it was developed. Four different versions of the CBCL (English version completed by 60% of the participants, the Chinese version 30%, Malay 8%, Tamil 2%) were used by Ooi et al. (2011), CARS in the USA and Spain while the DBC-ASA was used only in the USA. The SCQ was used in the UK, Qatar, Australia, and USA. The AABC was used only in Turkey. Out of the 12 screening tools for autism, the SCQ was used across a wider age range, across more disabilities, and comorbidities (Ung et al., 2016). The validity of the SCQ has also been examined in a small sample of children aged between 2.5 and 14-years in a South African study (Bozalek, 2013). When all assessment criteria for cross-cultural validity were examined, the overall rating for the autism tools was as follows: very low for the ADI-TSS, AQ, and MARA; low for the ASSQ, DBC-ASA, AABC and DiBAS-R; medium for EDUTEA, PDD-MRS, CBCL, CARS, and SCQ.

Out of the 19 studies reviewed for the intellectual disabilities screening tools, five studies used between groups designs, including samples of people without intellectual disabilities, six were within-subject designs, while the remaining eight were cross-sectional designs. Utilising the tool with different groups is a criterion for cross-cultural validity in the COSMIN, so studies that have not demonstrated this adequately were rated low in that area. HASI was used in two within-subject studies and five cross-sectional studies. SIT was used in three within-subject studies while the QT was used in one within-subject study. The LDSQ was employed in two between-

subject studies, the CAIDS-Q in 3 between-subject and one cross-sectional studies while the SCIL was used in two cross-sectional studies. The HASI was used in 4 different countries and languages: Norway, UK, Australia and Belgium. LDSQ was used in the UK and Scotland while the CAIDS-Q and QT were used in the UK and USA, respectively. The SCIL was used in Norway and the Netherlands. Putting together all the criteria for evaluating cross-cultural validity, the overall rating for the tools was moderate for CAIDS-Q and HASI, high for SCIL, low for the LDSQ and very low for both the SIT and QT.

Finally, given that one of the aims of this review was to consider the appropriateness of using these tools across a range of cultures, it is important to note that there are diverse cultures in Africa. These include a variety of spoken languages, beliefs and behaviours; therefore, whichever tools are identified through this review will require additional contextual adaptation and may perhaps benefit from further ethnological research.

4.7 Limitations

There are limitations to this review. By limiting the search to studies in English only, it is possible that some studies with adolescents, and potentially other tools, may have been missed. This in turn may limit the generalisability of the findings of this review, as there are some African countries whose official languages are not English. Manuals for some of the screening tools (seven in total) identified were not readily accessible. This meant that some information on validation reported in the studies could not be compared. Additionally, some of the administration and training requirements could not be examined in detail.

4.8 Conclusions

There are two main challenges as described in Chapter 2. The first relates to cultural adaptations and use of the tools outside of the development environment. Whichever tool is identified for use in Africa, it must be sensitive to local differences and expression. The language of the tool must be simple enough to understand, allowing for ease with translation or substitution where required. Validation of selected tools will require time, expertise and financial resources as determining the psychometric properties in a novel environment requires capacity. As such, the less complex the tool is, the easier it may be to assemble the required resources. These challenges are not to suggest the screening tools developed in the West are irrelevant to Africa or

LMICs, but that careful research and translational work may need to be done to ensure that a tool can be used successfully with people from other countries and cultures. A second challenge is that the tools selected for use with the adolescent and young adult population need to apply to a wide age range while remaining flexible and sensitive. Finally, the limited number of studies involving adolescents identified through this review has presented challenges (as in Hirota et al., 2018). Without a large body of knowledge about adolescents and continental Africa, particularly, the choice of tools is limited.

Developing and validating a continent-specific or country-specific tool for screening autism or intellectual disabilities will take considerable time, effort and resources. Such resources as time, training and personnel may not be readily available. Given the socio-economic and political climate of most African countries, the process could place a considerable financial burden on the economies. In summary, of the 18 tools (6 for intellectual disabilities and 12 for autism) identified through the review, except for the SCQ, none had been utilised in Africa. The SCQ was designed to be used with a wide age range, 4-years and above and has two versions (current and lifetime) which makes it a good fit for use with adolescents. The SCIL, meanwhile, was validated for adolescents and adults and includes test questions for intellectual abilities as well as social adaptive skills. The broad age range and inclusion of DSM-5 items places it above the other tools reviewed. Additionally, any tool that requires training and more than 20-minutes of administration time will add to the burden. Thus, to begin validating screening tools for autism and intellectual disabilities in African adolescents, the SCQ and the AQ seem appropriate for autism, while the SCIL and CAIDS-Q are for intellectual disabilities. The selection process for the best screening tool for both autism and intellectual disability is discussed in Chapter 5.

Chapter 5. Study 2 – Focus Group[2]

5.1 Introduction

Detecting developmental concerns suggestive of autism or intellectual disabilities using screening tools can help identify adolescents who need further diagnosis or intervention. Screening has been widely encouraged to identify persons with autism or intellectual disability, and many screening tools are available (Thabtah & Peebles, 2019). Consensus on the ideal and practical screening tools is lacking, however, more so where the tools are used in environments other than those in which they were developed (Marlow, Servili & Tomlinson, 2019). Ideally, screening for autism and intellectual disabilities should be part of children's routine visits to health professionals at an early age, but various factors hinder this in the Nigerian context. Some of these factors include the mindset of parents or caregivers and the lack of adequate resources (Franz, Chambers, von Isenburg & de Vries, 2017.) In practice, however, while opportunities for diagnosis may be missed in early childhood, adolescents are often identified when transitioning to secondary schools or facing more challenging environments and expectations of greater independence. In Africa, individuals with developmental disabilities are thus noticed either fairly late in schools or when parents seek medical attention for a severe illness (since visiting hospitals/health professionals for routine check-ups or minor ailments is not the norm), or at times when autism or intellectual disability specific research work is carried out (Knox et al., 2018; Saloojee, Phohole, Saloojee, & Ijsselmuiden, 2007; Gladstone et al., 2010; Scherzer, Chhagan, Kauchali & Susser, 2012; Bello-Mojeed, Omigbodun, Bakare, & Adewuya, 2017).

Well-developed screening tools for autism spectrum disorder or intellectual disability are readily available for younger children in the West and high-income countries (McKenzie, Paxton, Murray, Milanesi & Murray, 2012; Young, 2007; Robins, Fein, Barton & Green, 2001). Also, adaptations of existing screening tools for younger children have been conducted for other countries in the West (Nah, Young, Brewer, & Berlingeri, 2014; Canal-Bedia et al., 2011; Cuesta-Gómez, Manzone, & Posada-De-La-Paz, 2016; García-Primo et al., 2014). However, similar tools are not readily available for older children and adolescents, especially in low to

[2] This study is currently in production as, Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E. Using the consensus group method to select the best screening tools for autism and intellectual disability for use with Nigerian adolescents. Journal of Policy & Practice in Intellectual Disabilities.

middle-income countries. Very little work has been done in Africa and other low to middle-income economies regarding adapting existing tools for screening for either autism spectrum disorder or intellectual disability. Screening for autism and intellectual disability remains a challenge in low to middle-income countries such as Nigeria due to the absence of adequate tools and other factors such as denial and low level of awareness. Limited financial and human resources significantly contribute to the lack of adequate tools.

With the increasing global awareness of developmental disabilities such as autism spectrum disorder and intellectual disability (Malcolm-Smith et al., 2013), more individuals, especially younger children, now have early screening and intervention in the West. Indeed this is beginning to also happen in LMIC countries. Literature, however, remains extant on research involving older children and adolescents. In Nigeria, the significant challenges are with the older children and adolescents who have had no access to screening either by limited services or parents' choice. The lack of early identification leads to poor social integration, reduced quality of life and lack of intervention (Bargiela et al., 2016; Nwokolo, Langdon & Murphy, 2022). Addressing the adolescent screening gap requires a robust and culturally relevant measure with face and content validity. Resources in terms of financing and expertise are also potential barriers to developing new screening tools for low to middle-income economies; thus, adapting an existing tool is a prudent option. Substantial research on the adaptation of screening tools has been conducted in the West and other medium-income economies, where it is recognised that cultural disparities potentially impact adaptation (Long et al., 2020; Grinker et al., 2015). However, very little work has been done in Africa and other low to middle-income economies.

To begin addressing this challenge in countries such as Nigeria, the adaptation of existing screening tools should be considered. Adapting existing tools is the most common and fastest approach to creating usable screening tools for countries with limited resources or expertise. However, concerns have been raised about the feasibility of employing adapted tools for screening across cultural groups (Soto et al., 2015). One way of addressing these concerns is to follow clearly defined methodologies such as those stipulated by the International Test Commission (International Test Commission, 2017). Pertinent aspects of the methodology include examining the tool's content validity, cultural validity, and language by engaging the relevant experts. Cultural validity assesses whether constructs and language initially generated in

a single culture are appropriate, relevant, applicable, equivalent, and meaningful in another culture (Matsumoto & Yoo, 2006; Beaton et al., 2000). Content validity, which ensures that the items in a screening tool represent all relevant aspects of a given construct, is one of the essential psychometric properties of a screening tool (Mokkink et al., 2018, Prinsen et al., 2018; Terwee et al., 2018a). Cultural and content validity outside of the environment where the tool was initially developed is usually examined by a group of experts in the environment concerned, in this case, Nigeria (International Test Commission, 2017).

During the completion of a recent systematic review (Nwokolo et al., 2022), twelve (12) screening tools for autism spectrum disorder (ASD) and six (6) for intellectual disability (ID) were identified. Of these, four tools were chosen (two tools each for ASD and ID) for use within the current study based on the cross-cultural validity and overall quality ratings of studies developing the tools. The two tools for ID were (a) the Screener for Intelligence and Learning Disabilities (SCIL) (Nijman et al., 2018), a standardised 14-item questionnaire developed and used for adolescents in the Netherlands, and (b) the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q) (McKenzie & Paxton, 2012), a short 7-item screening questionnaire. The SCIL was originally in Dutch, and as part of this study, translated to English, while the CAIDS-Q was in English. For screening ASD, the measures selected were (a) the Social Communication Questionnaire (SCQ) (Rutter et al., 2003) and (b) the Autism Spectrum Quotient (AQ-10), adolescent version (Allison et al., 2012). To adapt any screening tools for use in Nigeria, selecting suitable and culturally sensitive measures was crucial. Thus, a consensus group of the relevant professionals and lay people resident in Nigeria were recruited for the study. The aims of the study were to consider the face, content, and cultural validities of our chosen screening tools and make recommended adaptations for use with Nigerian adolescents using a consensus group methodology.

5.2 Methods

5.2.1 Consensus Method

The consensus group methodology was chosen due to its extensive use in studies for similar decision-making (Humphrey-Murto et al., 2017; International Test Commission, 2017). The process is based on the notion that valid, accurate and reliable evaluation can be best achieved by

consulting a team of experts and stakeholders. Achieving accurate and reliable assessment is assumed achievable through the group (Humphrey-Murto et al., 2017). Consensus methods have been used in education for curriculum development (O'Neil & Jackson, 1983), in medical research (Humphrey-Murto et al., 2017) and in health studies for planning (Van de Ven & Delbecq, 1972). Studies supporting the use of the consensus group methods in developing items for measurement tools, developing clinical guidelines, and deciding on components of new or revised curriculums all exist (Van de Ven & Delbecq, 1972; Murphy et al., 1998; Humphrey-Murto et al., 2017). For instance, O'Neil & Jackson (1983) used the Nominal Group Technique (NGT) to assess the contents of an existing third-year university course, while Van de Ven & Delbecq (1972) used the NGT to ascertain the qualitative dimensions of a comprehensive healthcare program. In these studies, the consensus groups examined, for example, the relevance of healthcare consumers, the language of the tools and the content of the existing coursework. These are similar to the intent of using the consensus group in assessing the identified autism and intellectual disability screening tools in this study. Another reason for using consensus methods is that they control for possible researcher bias. An appropriate and systematic process must be employed to select the best option, outcome, or measure. Using the consensus group method has been shown to be such a technique (Delbecq, 1967; Hutchings et al., 2010 & 2012). Consensus methods are considered qualitative and a systematic means for determining and developing consensus. The goal is to establish how well experts and stakeholders agree on an issue through consultation and accepting the group agreement (Tammela, 2013). This method also allowed for a consideration of the cultural relevance of each measure and for associated adaptations to address any issues.

Two main techniques are used for consensus group meetings: the Nominal Group Technique (NGT) or the Delphi method. With each method, questions are raised, solutions are proffered, and responses are ranked and agreed upon. Each of these methods has its strengths and weaknesses. Although the Delphi method is used quite often for the development of initial research questions and involves a large number of participants who are anonymous, the Delphi method limits discussions. The NGT, on the other hand, involves a smaller number of participants, it allows for face-to-face discussions and debates. Given that we chose existing tools that had been previously developed and we aimed to ascertain the cultural relevance, the NGT was chosen as it allows for extensive discussion.

The nominal group technique was used to review, evaluate, and consider our screening tools' face, content, and cultural validities within a Nigerian context and make any associated adaptations. The technique has also been applied for problem-solving and planning (Delbecq & Van de Ven, 1971), team decision-making (Bartunek & Murninghan, 1984) and as a research instrument (Van de Ven & Delbecq, 1972). NGT is a semi-quantitative, highly structured and facilitated group-based decision-making process. The process is deemed an excellent form of brainstorming with limited member-to-member discussions. Facilitation of discussions allows for and encourages the active participation of all members and prevents the potential of an individual member's dominance of the discussions (McMillan et al., 2016; Murphy et al., 1998). The face-to-face interactive nature of the NGT usually involves 5 – 12 participants (O'Neil & Jackson, 1983; Tammela, 2013; Humphrey-Murto et al., 2017). Where the group size is greater than this, the suggestion is that sub-groups of 8 – 10 members can be formed (O'Neil & Jackson, 1983). Delbecq and Van de Ven (1971) outlined the nominal group model's implementation process. Although there is a set of guidelines and a structure for using the NGT, in practice, the techniques have been varied based on the project or user requirements (McMillan et al., 2016; Murphy et al., 1998). Such variations may be due to the participants' time, level of clarification, research goals, or consensus. At other times, the requirement may be an adaptation to the stages, such as reviewing an existing protocol, measure or where the population is culturally or linguistically diverse (McMillan et al., 2016). A modified NGT was used here to select and decide which autism and intellectual disability screening tools would be used for the validation study.

5.2.2 Choice of Experts

Experts, in the context of the NGT, are individuals who are knowledgeable about the subject matter. Given this objective, the recruitment of experts – psychologist, speech pathologist, behavioural technician, psychiatrist, teacher and paediatrician – was purposive to include members from the relevant professions with professional experience and knowledge of the relevant population. For existing measures, content validity is evaluated by systematically asking professionals and users about the comprehensiveness, relevance, and comprehensibility of the items (Terwee et al., 2018b). A parent and layperson were also included in order to assess the comprehensibility of the screening tools, while comprehensiveness and relevance were assessed

by the professionals (Terwee et al., 2018b). Inclusion of the layperson and parent in the group was based on the different benefits outlined by Delbecq and Van de Ven (1971) and Van de Ven and Delbecq (1972). First, it eliminated the sole focus on the professional perspective. Secondly, the user's needs and perspective, in this case, the parent's, are included and finally, it allows a more robust assessment of the screening tools because of the user's participation and representation in the decision-making process. This professional and public group method was previously utilised in several health-related studies (McMillan et al., 2015; Tammela, 2013).

Through the main researcher's networks, experts were either identified through parent networks or recommended by general practitioners who were approached and asked to share information about the study. Seventeen experts, parents, and laypersons were invited via email, telephone messages and personal contact. Participants were given three possible meeting dates and asked to provide feedback on availability. They were followed up via email, telephone calls and chat messages, with several reminders sent to the non-responders. Following telephone and chat responses, the proposed meeting dates and schedules were shared with eight individuals who confirmed their availability. All were provided with the information sheets about the study.

5.2.3 Participants

The participants comprised a group of 8 individuals, of whom 60% were female, and 40% were male. The participants' ages ranged between mid-thirties to mid-fifties, and all were middle-income urban dwellers. The group consisted of a psychologist, a psychiatrist, a teacher, a paediatrician, a behavioural technician, a speech pathologist, a layperson with a background in information technology and a parent (Table 10).

Table 10 - Distribution of participants

Profession	Sex	Age Bracket
Psychologist	Male	Late 40s
Psychiatrist	Male	Late 40s
Paediatrician	Male	Late 50s
Teacher	Female	Late 50s
Parent	Female	Mid 30s
Behaviour Technician	Female	Mid 30s
Layperson	Female	Late 30s
Speech Pathologist	Female	Late 40s

5.2.4 The Meeting

The meeting started late morning and lasted six hours with a one-hour lunch break. The researcher, who also facilitated, made a 15-minute presentation to provide background information on the project and a summary of the systematic review results (Nwokolo et al., 2022) for the participants. After that, the nominal group process was explained, and the participants were given the consent form to read and sign. Consent included granting permission to record the meeting. The participants were assured that all information would be anonymised and treated confidentially. Signed consent forms indicated a willingness to participate. Ethical approval for the study was obtained from the University of Kent, Tizard Centre Ethics Committee, and the National Health Research Ethics Committee of Nigeria (NHREC; NHREC/01/01/2007-16/09/2019).

Additionally, the researcher explained the goal and expected outcomes to the participants. Once all questions were answered and clarity provided, the screening tools to be reviewed were handed out. The meeting was organised in two sessions: the first segment discussed the autism tools, while the intellectual disability tools were discussed in the second half of the session. As the screening tools were not redesigned, the NGT method was modified (McMillan et al., 2016). Phases one and two were merged, and the first step – silent generation of ideas - was modified to review each measure's existing format, questions, and content. After that, one measure was selected for autism and one for intellectual disability and reviewed in detail. Phases three to five

were merged for the second stage. During the second stage, the discussion was open, and group members' interactions were allowed but moderated by the facilitator. Allowing open discussion and interaction was a culture-based decision that had minimal influence on the individual suggestions and conclusions. During the discussions, ideas and comments were stated in a round-robin manner, with clarifications given. The facilitator collated all suggestions, votes, and agreements. The entire meeting was recorded, transcribed, and analysed for themes.

5.2.5 Measures

5.2.5.1 Autism Spectrum Disorder Screening Tools

The two screening tools reviewed were the Social Communication Questionnaire (SCQ) and the Autism Screening Quotient (AQ-10) adolescent version. Both measures were identified via a systematic review (Nwokolo et al., 2022). The SCQ is a brief 40-item parent or caregiver-report screening measure used widely in research (Berument et al., 1999). The measure has two versions, the lifetime version and the current version, both focusing on symptoms of autism most likely to be observed by the individual's principal caregiver. The SCQ is designed for anyone 4 years old and above, and it takes about 10 – 15 minutes to complete and about 5 minutes to score. The lifetime version was used in this study, given the intended age range of participants (11 – 26 years) for the validation study. In addition, Wei, Chesnut, Barnard-Brak, & Richman (2015) reported that the lifetime version had better psychometric properties than the current version. The AQ-10 is the short version of the AQ-50 and is usually completed by a parent or caregiver. The AQ-10 adolescent version can be completed in 10 minutes or less and was found to have good psychometric properties based on the systematic review (Nwokolo et al., 2022). Also, it is adolescent-specific, which is useful, given the age range of the participants for the validation study. The lifetime version of the SCQ and the AQ-10 were presented to the consensus group participants.

5.2.5.2 Intellectual Disability Screening Tools

Two tools identified through a systematic review were presented to the participants (Nwokolo et al., 2022). The tools were the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q) and the Screener for Intelligence and Learning Disabilities (SCIL). The CAIDS-Q is a short 7-item screening questionnaire for detecting intellectual disabilities in

children and adolescents developed by McKenzie and Paxton in 2012. The SCIL was developed as a 14-item screening tool in the Dutch language (Nijman et al., 2018; Geijsen et al., 2018). There is no commercially available English version of the SCIL. Translation from Dutch to English was therefore performed, following the procedure laid out by the International Test Commission (International Test Commission, 2017). To ensure that the overlap in definition and constructs measured were adequately captured, a 2-person expert and bi-lingual team of clinical psychologists in the field of intellectual disability translated the Dutch version to English. Both team members were Dutch; one was resident in the United Kingdom, and the other in the Netherlands. English-only speaking clinical psychologists reviewed the English version. The English translation was then sent back to the Dutch developers to be translated back into Dutch and finally back into English. Internationally, the back translation and adaptation process is often used to ensure that linguistic equivalence, psychological and cultural differences are considered (Grisay, 2003; International Test Commission, 2017). Usually, the source version (Dutch) of the text is translated into the intended version (English) and then translated back to the original language for comparison and identification of possible discrepancies. This back-translation technique is useful for detecting essential interpretation issues or mistranslations (Hambleton, 2002; Grisay, 2003). Once the English version correctly reflected the Dutch version's content, structure, and language, the research team finalised the arrangement and utilised it with the Nominal Group. Both measures (SCIL and CAIDS-Q) were designed for use with adolescents, have good psychometric properties, and have been used in various studies (Nwokolo et al., 2022). Also, given the age range (11 – 26 years) of the intended participants in the Nigerian validation study, both tools were deemed appropriate.

5.3 Procedure

Ethical approval for the study was obtained from the University of Kent, Tizard Centre Ethics Committee, and the National Health Research Ethics Committee of Nigeria (NHREC; NHREC/01/01/2007-16/09/2019). Copies are in the Appendix (Appendix 1 and 3).

The consensus meeting started late morning and lasted six hours with a one-hour lunch break. The researcher, who also facilitated, made a 15-minute presentation to provide background information on the project and a summary of the results of the systematic review (Nwokolo et al., 2022) for the participants. Following that, the nominal group process was explained, and the

participants were given the consent form (Appendix 5 & 6) to read and sign. Consent included granting permission to record the meeting. The participants were assured that all information would be anonymised and treated confidentially. Signed consent forms indicated a willingness to participate.

Additionally, the researcher explained the goal and expected outcomes to the participants. Once all questions were answered and clarity provided, the screening tools to be reviewed were handed out. The meeting was organised in two sessions: the first segment discussed the autism tools, while the intellectual disability tools were discussed in the second half of the session. As the screening tools were not redesigned, the NGT method was modified (McMillan et al., 2016). Phases one (problem exploration) and two (knowledge exploration) were merged, and the first step – silent generation of ideas - was modified to review each measure's existing format, questions, and content. After that, one measure was selected for autism and one for intellectual disability and reviewed in detail. Phases three to five (priority development, program development and program evaluation) were merged for the second stage. During the second stage, the discussion was open, and group members' interactions were allowed but moderated by the facilitator. Allowing open discussion and interaction was a culture-based decision that had minimal influence on the individual suggestions and conclusions. During the discussions, ideas and comments were stated in a round-robin manner (one participant at a time stated a single idea to the group), with clarifications given. The facilitator collated all suggestions, votes, and agreements. The entire meeting was recorded, transcribed, and analysed for themes.

5.3.1 Measure Review

The participants were given the four screening tools (two each for autism spectrum disorder and intellectual disability), the SCQ, AQ-10, CAIDS-Q and SCIL, to review. Participants were asked to assess the face validity, content validity and cultural relevance of all four tools. To assess face validity, the participants were required to evaluate the items with respect to language, ambiguity, interpretability, comprehensibility, understandability, and familiarity of items (Mousazadeh, Rakhshan, & Mohammadi, 2017; Mokkink et al., 2018). For content validity, all participants except the parent and layperson assessed the comprehensiveness, applicability in practice, understandability, and relevance to the Nigerian context. The open discussion allowed the parent

and layperson to flag potential challenges that may be encountered in practice. Following the assessment, a comparison was made between the SCQ, and the AQ-10 and the pros and cons were discussed. Similarly, the group compared the SCIL to the CAIDS-Q. In-depth discussion of the preferred measures followed with the facilitator's guidance. Ambiguous words and examples were clarified, and more culturally relevant words or phrases were suggested. After the discussion and clarification, the suggested options were voted on and selected.

5.3.2 Data Analysis

5.3.2.1 Consensus

Although Fink, Kosecoff, Chassin & Brook (1984) stated that there are no specific rules for establishing consensus, they describe the various criteria, such as percentage of participants in support, topics with the most votes, and rating on a scale. Fink et al. (1984) also mentioned that the narrower the criteria, the more challenging obtaining consensus usually is. Given that consensus meetings aim to determine the extent of agreement between experts, the threshold for agreement is typically predetermined. Williamson et al. (2012) and Humphrey-Murto et al. (2017) suggested that advance consideration and a clear definition be given to the criteria for consensus. Various thresholds have been reported in the literature as acceptable; 67% (Cantrill, Sibbald & Buetow, 1996), 75% and 80% (McConachie et al., 2018), while Williamson et al. (2012) suggested 70% for consensus. The extent to which each participant agrees with the contents of each measure under consideration was defined as agreement. Based on Williamson et al. (2012), a criterion of 75% threshold was set for this study. The threshold of 75% meant 6 out of the 8 participants (Fink et al., 1984) had to agree on the retention of the original wording of the measure or with the suggested modification. A simple response tallying for each question was used, and percentage agreement was calculated. For the SCQ, each of the 40 questions was analysed separately and similarly for the 14 questions of the SCIL 14 – 17. All data were collated and analysed using Microsoft Excel for Windows 10.

5.3.2.2 Meeting Transcription and Theme Generation

Because consensus methods are considered to be qualitative methods (Tammela, 2013; Jones & Hunter, 1995), the meeting recording was transcribed and analysed following the thematic analysis (TA) methodology (Braun & Clarke, 2006; Alhojailan, 2012). Thematic analysis has

been used before to analyse NGT data (McMillan et al., 2014; Søndergaard et al., 2018). To mitigate against eclipsing individual positions, individual idiosyncrasies are included as themes, with the reverse also being applicable, where the group is not eclipsed while privileging the individual. For this study, a combination of the process and modifications outlined in Tomkins & Eatough (2010) and Palmer et al. (2010) were used. Tomkins & Eatough (2010) employed a superordinate (individual level) theme analysis while maintaining the group interactive context. Palmer et al. (2010) explored the participants' experiential claims and concerns followed by a development of a parallel commentary in the context of the group discussion.

The following steps were implemented in analysing the data with an explanation of what was done.

1. Familiarisation with the data. Familiarisation involved the first author transcribing the data and re-reading the transcript at least three times while appraising each participant's comment and contribution. Noting of initial ideas also occurred.
2. Initial codes generated. Codes were generated based on the meaning of each participant's thoughts and were colour coded. Comments were made on the right-hand side of the margin about the meaning.
3. Searching for themes. The colour-coded texts were clustered into potential themes on a group level. Coloured words, phrases, sentences, and passages were re-read to get a sense of the overall perspective from a particular participant without eclipsing the group. Each colour represented an emerging theme.
4. Collating codes into themes. All data were extracted and gathered into relevant main and sub-themes. Main and sub-themes were produced and named (Table 11). These themes are described in some detail with reference to direct quotes from the participants.
5. Reviewing themes. Themes were cross-checked relative to the codes with ongoing analysis to refine the specifics of each theme and the overall story.
6. Producing the report. Examples of effective extracts were selected and analysed for inclusion in the study report. The selection of the extracts was made relative to the research question.

Data trustworthiness is relevant in qualitative research work and Nowell, Norris, White & Moules (2017) outlined the process to ensure data trustworthiness. The process expands on the steps outlined in Braun & Clarke (2006). Trustworthiness is measured by credibility, transferability, dependability, and confirmability criteria. In phase 1, for instance, the process requires prolonged engagement with the data and maintaining records of all data field notes and transcripts. The data were reviewed thrice during transcription, with continuous reference to the data while putting together the study report. The raw data and original notes were stored in a secure place. Other steps suggested by Nowell et al. (2017) are team consensus on themes in phase 5, member checking (phase 6) and documentation of meetings (phase 2). The research team members (second & third authors) vetted the themes and sub-themes proposed by the first author and reached an agreement. Additionally, the summary of all meeting sessions was documented and stored via a secure system.

5.4 Results

5.4.1 Meeting Outcome

Seven out of the eight participants were present at the start of the meeting. The eighth participant joined about 40 minutes later—another participant left due to a prior engagement about an hour before the end of the meeting. Although one participant joined late and another exited early, the agreement calculation was based on the total number of participants, eight. However, this had no significant impact on the results reported in the relevant sections below, as the threshold of 75% agreement set for the study was exceeded (details are in sections 5.4.2.2.1 & 5.4.2.2.2).

For screening intellectual disability, the participants chose the SCIL 14 – 17 as they found it more robust and thorough, stating that they felt the CAIDS-Q was overly simplified. The group indicated that the SCIL 14 – 17 tested the relevant skills such as intellectual functioning and some adaptive skills. Similarly, for screening autism spectrum disorder, the SCQ was chosen over the AQ-10 as more robust and comprehensive with questions that examine the relevant autism spectrum domains.

5.4.2 Themes

Following the analysis of the transcript, three themes were identified. Namely language, cultural relevance, and face validity. These are listed in Table 11.

5.4.2.1 Language

This theme focuses on how Nigerians use the English language and the meaning attached to certain words, sometimes depending on the context. The word '*rituals*' used in question 8 of the SCQ was deemed to have a negative connotation, and the participants advised that an alternative word be used. In the African context and Nigeria, rituals involve sacrifices to 'deities' or some god. The word 'rituals' was therefore changed to 'routines'. Meanwhile, Question 9, 'has her/his facial expression usually seemed appropriate to the particular situation, as far as you could tell?' on the SCQ elicited the following dialogue:

> *R: how do we determine what appropriate facial expression is?*
>
> *BK: to the situation, it says 'to the particular situation'. For instance, someone is dead, and you're smiling.*
>
> *AB: or they're supposed to be afraid or scared*
>
> *AO: again, one of the things I have come to realise is that there is a Nigerian English. If I want to say that thing, I may say that 'has her/his facial expression often reflected the situation at hand', as far you could tell?*

A good number of Nigerian dialects are spoken with a double emphasis, which may appear as either verbal or logical tautology when translated to English. In Yoruba, for instance, the phrase '*pada sẹhin*', when translated to English, means '*return back*'. Thus, AO stated that 'there is a Nigerian English'. Another example was item 14 of the SCQ, 'has she/he ever seemed to be *unusually* interested in the sight, feel, sound, taste, or smell of things or people?'. The discussion was as follows:

> *AB: sorry to take you back to #14. Even though it cuts across all senses, some persons, when you talk about feeling things, may not be able to relate that to touch. So how do you go about that?'.*
>
> *Me: the parent or the individual…?*
>
> *BK: what I hear him say is that the word 'feel' in this context may be interpreted emotionally as opposed to tactile*

AB: is there a way to put 'touch' in brackets?

All: tactile

GB: that one is grammar

OO: 'touch' is more appropriate for our environment than tactile

R: 'tactile' sounds really oyibo, 'touch'

AO: there is Nigerian grammar even with academic papers. The editor will ask you to find a native English speaker who will edit, who knows exactly what you are saying but puts it in a different way. But when you are dealing with instruments like this, I believe the more you 'Nigerianise' it, the more you'll get the appropriate response

Further discussion considered Question 31 of the SCQ, which asks, 'when she/he was 4 to 5, did she/he ever try to comfort you if you were sad or hurt?'. Since Nigerians say 'sorry' for nearly every incident, including those the individual is not responsible for, the group recommended adding examples for clarity.

For the SCIL 14 – 17, language reference was minimal. The agreement was to change the word 'GP' to 'doctor' as the term 'GP' is not used in Nigeria. Regarding the dictation component of the measure in question 12, the group agreed to exclude words with consonants likely to be mispronounced to avoid possible h-dropping (such as hitting). Question 13 of the SCIL 14 – 17 tests reading skills, and the ability to read fluently incorporates the reader's comprehension, familiarity with the words and background knowledge of the context. Question 13 in the SCIL 14 – 17 includes these words: '*pay for parking by mobile phone. When you have parked your car, log in on your mobile using the (location) code as advertised/displayed on the signs and parking machines. When you leave, you log out by phone/mobile*.' In the Nigerian environment, parking is not paid for like this. However, an approximate equivalent is the point of sale (POS) machines with bank cards used in stores. To use language that will be familiar in the Nigerian context, the group agreed on 'bank card'. Below are some excerpts from the discussion.

BK: this is based on those places where you have parking metres. Then you slot in and pay for your parking. Where there's no context for it…

AO: parking at the mall. No, you don't even need to do the possibility. Just say, the process of…

R: are you allowed to change it completely?

BK: no, you're turning it into a title. It's a sentence. It tells you it is possible to pay by phone, then it now telling you how to do it.

Me: (tell my story). In this context, in order not to change the story completely, we can say "ATM" or "POS".

R: can we say 'card'? Is it everybody that knows 'POS'?

BK: yes, is it not every Nigerian that knows "POS"?

R: adolescents?

BK: yeah, it's the language of the environment.

AO: yes, it is. "POS" is the language, but I don't want us to introduce a word that is not actually a word; "POS".

Me: ok, so, with 'card' because 'POS' is 'point of sale'. So, by 'card'.

Table 11 - Main themes and sub-themes for the SCQ and SCIL 14-17)

Main Theme	**Sub-themes**
Language	• Use of words • Meaning of the word • Context • Nigerian parlance
Cultural relevance	• Examples given • Family dynamics (the way parents relate with their children) • Context
Face Validity	• A professional versus the parent's understanding of the question • Environment

5.4.2.2 Cultural Relevance

There were three sub-themes under cultural relevance: the examples given, family dynamics (the way parents related with their children), and context. All the examples given related predominantly to the design environment, the West. The group advised utilising more culturally

relevant examples. For instance, vacuuming, gardening, or mending things were given as examples in question 21 on the SCQ. In the Nigerian context, not everyone vacuums, and mending things appeared vague. Therefore, the participants suggested using examples such as sweeping and washing. A portion of the dialogue follows below.

AO: sweeping, more people sweep than they vacuum even if they are cosmopolitan or whatever group we are looking at

BK: maybe just cleaning, washing

OO: that's appropriate. Just look at things that we do here

GB: local content

Following the discussion on question 21, the group agreed that the questions were relevant and appropriate from questions 32 and below. However, for some questions, more local examples and songs were suggested as replacements of Western ones. GB mentioned activities such as 'backing a baby' (a traditional African method where mothers carry babies and infants on their backs swathed in cloth), 'cooking with hibiscus flower', 'playing mummy and daddy'. At the same time, AO said, *"I see that even in real practice, what differentiates what we do at times from questionnaires alone, is that opportunity to spend time explaining what we do, unlike just giving it to them to fill. You realise that the more you are engaging, the more the individual is able to know exactly what you are talking about."* Buttressing AO's point, AB said, *"which is what I've found with parents most often. When you give them a questionnaire like this, what they do is to fill, and when they get to where they don't understand, they will ask a question. Once you give them examples, it's clear, and they give you other examples."* AO, *"so meaning that a useful questionnaire in this environment will do well to have short-short examples where necessary, which is what we are doing."*

Turning to family dynamics, question 2 on the SCQ designed to assess the extent of vocalisation may require some explaining as holding 'to and fro conversation' is not the norm in typical Nigerian homes. Although there is some shift regarding this, children are often expected to respond to questions asked by parents, rather than engage in 'chit chat'. The younger parents are at the fore of changing this narrative. One of the younger participants, AO, said, *"to and fro, they may get a little bit but once you say converse (pause), in fact, a lot of people complain that they are coming to come and tell you that yeah, they are talking, but he is still having problems with*

conversation." Therefore, the group agreed to leave the question as is and give examples of what a 'to and fro conversation' entails.

5.4.2.2 Face Validity

The last theme, face validity, covers environment and professional versus parent's understanding of a question. As the SCQ is a self-administered (parent) questionnaire, the participants opined that it might be more useful if the professionals administered it to allow for explanations where there is the possibility of confusion or lack of clarity. For instance, question 4 reads as *'has she/he ever used socially inappropriate questions or statements? For example, has she/he ever regularly asked personal questions or made personal comments at awkward times?'* To which the following dialogue ensued.

> *GB: when a child is done eating, there is no need to say, 'will the food be ready'.*
> *AO: I'm thinking that while I agree that it is clear, we must also remember that if you are very familiar with ASD, some of these questions will be clear to you. But if you are not familiar with ASD, you may not actually grasp it. This particular question, we all know what this question is trying to test.*
> *Me: that's why I'm looking at my parent; as a parent, if you are given this question, is this clear enough. Are you able to answer yes, or no?*
> *R: yes, but I am a parent who already knows quite a bit. Going back to what he's saying, I am not a lay parent that has just come.*
> *BK: the idea of inappropriateness, from the example given, it's one more out of context versus something more socially inappropriate in terms of asking a personal question.*

Once all participants had expressed their opinions, the group agreed that the correct response, 'yes' or 'no', would be elicited from respondents irrespective of their background.

Regarding the screening tool for intellectual disability, the SCIL 14 – 17, the group discussed questions 1 and 2 extensively. The questions are centred on special education and level of education. Many Nigerian schools in the urban areas purportedly offer special education services.

> *OO: looking at question 1 for me, looks like the first stem and second stem are looking at the same thing.*

GB: but in the true sense of it for people practising, for example, 'did you receive special education?' You can be in a regular school system and be receiving support from a unit.
PA: yes, and you're receiving support from a unit. Yes.
GB: do you go to a special needs school? You could have a school that is a special needs school, all the teachers there are specialist trained personnel, and you have special materials, and that school is labelled for that specific learning difficulty. It may be school for hearing impaired, school for individuals with learning disability or school for individuals with autism. So, you could have that, or did you have a special education need? That means are you having challenges with learning, typically. So, the three questions are not actually the same. We could sample different people differently.
AO: in any case, the answer is 'yes' or 'no'. Meaning that when you read through the question, anyone of it is what you are responding to.
GB: you will fall into one category. The one that applies to you.

Once the different educational categories and services were agreed on, the team accepted the questions. The levels of education were also expanded to include the different curricula, both national and international, offered in the country. Some of these are the West African Examination Council (WAEC), polytechnic, monothecnic and teachers' colleges.

In discussing question 3, *'do you receive or have you received support from a service for people with Intellectual Disability (excluding a home tutor or lesson teacher)?*' was examined at length by the participants. The exclusion of home tutors and lesson teachers was the consensus as there are no such services for a person with intellectual disability in Nigeria. A private tutor (lesson teacher) is typically employed once a child struggles in school. However, some who do not struggle with schoolwork have these tutors as a competitive advantage. The distinction lies in their academic performances, so having a private tutor does not necessarily indicate a pupil is struggling to understand the material.

For both the SCQ and the SCIL 14 – 17, the consensus from the group was that face validity was met. The items on the SCIL 14 – 17 have specific and relevant questions that test for intellectual disability. At the same time, the different DSM-5 domains (social and communication deficits, repetitive and restricted behaviours) of autism spectrum disorder are captured in the SCQ.

5.4.2.2.1 Social Communication Questionnaire (SCQ) for screening Autism Spectrum Disorder

The participants agreed that several changes were needed, such as that more local and culturally relevant examples should be given in the tool. For instance, for question 6 ("has she/he ever used words that she/he seemed to have invented or made up her/himself; put things in odd, indirect ways; or used metaphorical ways of saying things (e.g., saying *hot rain* for *steam*)?"), experts' opinion was to give examples to the respondents in context with Nigeria. Thus, for question 6, the replacement for "*hot rain for steam*" would be "*jagbajantis for mess*".

Another example was question 8; the word 'ritual' was explained as 'routine' to remove any fetish connotation. According to Hambleton (1996, p. 28), "when an instrument is adapted for use in another population, documentation of the changes should be provided, along with evidence of the equivalence." The list of examples of other culturally relevant words, examples, and clarifications are in Table 12. Overall, the participants agreed that 23 (58%) of the 40 items were culturally relevant and required no modification. After discussions and adaptations, between 87.5% and 100% agreement were achieved for all 40 questions. Table 12 shows the SCQ questions which were modified.

Table 12 - List of questions and the agreed cultural examples and modifications for the SCQ

SCQ Item number	Number of votes	Comments and suggested clarifications
1	7	Include examples such that it is clearer (mummy see, etcetera. # of words).
6	8	Give examples of respondents in context Nigeria, e.g., 'jagbajantis' for mess
8	8	For the word 'rituals' use "routine"
9	8	Example laughing when something is funny or showing concern when something is wrong
12	8	Include a 2nd example - combing the doll's hair over and over, switching a torch on and off
13	8	Examples are male dominant; add dolls, etcetera. For females
14	8	For feel put 'touch' in brackets

15	8	Include ‘face’’
16	8	Examples - hanging upside from a chair, twisting their body into a funny shape, any unusual body movement
18	8	Give other examples - cars, dolls, something that seems like a favourite item
20	8	The words in bracket meant for clarification (‘rather than to get something’), we can use “only to get something”
21	8	Local examples such as sweeping, cleaning the table, washing plates
28	8	For engage, put “get” & “keep” in brackets
30	8	Add, e.g., playing hide and seek
31	8	Add e.g., “say sorry”
33	8	Example in brackets (sad, etc.)
34	8	Examples of local songs and common ones; “if you’re happy”, “ABCD…”, “twinkle twinkle”, “xxx is a good girl or boy”
35	8	Example playing daddy & mummy, backing a baby*

**a traditional African method where mothers carry babies and infants on their backs swathed in cloth*

Table 13 - List of old questions and their modifications for the SCQ

SCQ Item number	**Old question**	**Modified question**
1	Is she/he able to talk using short phrases or sentences? If *no,* skip to question 8.	Is she/he able to talk using short phrases or sentences? If *no,* skip to question 8. How many words can she/he use when talking? For example, ‘mummy see’, ‘come here’, ‘what is your name?’
6	Has she/he ever used words that she/he seemed to have invented or made up her/himself; put things in odd, indirect ways; or used metaphorical ways of saying things (e.g., saying *hot rain* for steam)?	Has she/he ever used words that she/he seemed to have invented or made up her/himself; put things in odd, indirect ways; or used metaphorical ways of saying things (e.g., saying *jagbajantis* for mess)?
8	Has she/he ever had things that she/he seemed to have to do in a very particular way or order or rituals that she/he insisted that you go through?	Has she/he ever had things that she/he seemed to have to do in a very particular way or order or routines that she/he insisted that you go through?

9	Has her/his facial expression usually seemed appropriate to the particular situation, as far as you can tell?	Has her/his facial expression usually seemed appropriate to the particular situation, as far as you can tell? For example, laughing when something is funny or showing concern when something is wrong.
12	Has she/he ever seemed to be more interested in parts of a toy or an object (e.g., spinning the wheels of a car), rather than using the object as it was intended?	Has she/he ever seemed to be more interested in parts of a toy or an object (e.g., spinning the wheels of a car, combing the doll's hair over and over, switching a torch on and off), rather than using the object as it was intended?
13	Has she/he ever had any special interests that were unusual in their intensity but otherwise appropriate for her/his age and peer group (e.g., trains, dinosaurs)?	Has she/he ever had any special interests that were unusual in their intensity but otherwise appropriate for her/his age and peer group (e.g., trains, dinosaurs, dolls, clothes)?
14	Has she/he ever seemed to be unusually interested in the sight, feel, sound, taste, or smell of things or people?	Has she/he ever seemed to be unusually interested in the sight, feel (touch), sound, taste, or smell of things or people?
15	Has she/he ever had any mannerisms or odd ways of moving her/his hands or fingers, such as flapping or moving her/his fingers in front of her/his eyes?	Has she/he ever had any mannerisms or odd ways of moving her/his hands or fingers, such as flapping or moving her/his fingers in front of her/his eyes or face?
16	Has she/he ever had any complicated movements of her/his whole body, such as spinning or repeatedly bouncing up and down?	Has she/he ever had any complicated movements of her/his whole body, such as spinning, repeatedly bouncing up and down, hanging upside from a chair, twisting their body into a funny shape, any unusual body movement?
18	Has she/he ever had any objects (other than a soft toy or comfort blanket) that she/he had to carry around?	Has she/he ever had any objects (other than a cars, dolls, something that seems like a favourite item) that she/he had to carry around?
20	When she/he was 4 to 5, did she/he ever talk with you just to be friendly (rather than to get something)?	When she/he was 4 to 5, did she/he ever talk with you just to be friendly (rather than only to get something)?
21	When she/he was 4 to 5, did she/he ever *spontaneously* copy you (or other people) or what you were doing (such	When she/he was 4 to 5, did she/he ever *spontaneously* copy you (or other people) or what you were doing (such as sweeping, cleaning the table, washing plates)?

	as vacuuming, gardening, or mending things)?	
28	When she/he was 4 to 5, did she/he ever show you things that interested her/him to engage your attention?	When she/he was 4 to 5, did she/he ever show you things that interested her/him to engage (get & keep) your attention?
30	When she/he was 4 to 5, did she/he ever seem to want you to join in her/his enjoyment of something?	When she/he was 4 to 5, did she/he ever seem to want you to join in her/his enjoyment of something (e.g., playing hide and seek)?
31	When she/he was 4 to 5, did she/he ever try to comfort you if you were sad or hurt?	When she/he was 4 to 5, did she/he ever try to comfort you if you were sad or hurt (e.g., say sorry)?
33	When she/he was 4 to 5, did she/he show normal range of facial expressions?	When she/he was 4 to 5, did she/he show normal range of facial expressions (e.g., sad, angry, happy etc.)?
34	When she/he was 4 to 5, did she/he ever spontaneously join in and try to copy the actions in social games, such as *The Mulberry Bush* or *London Bridge Is Falling Down*?	When she/he was 4 to 5, did she/he ever spontaneously join in and try to copy the actions in social games, such as *ABCD, Twinkle Twinkle Little Star, If You're Happy and You Know It Clap Your Hands*, or *XXX is a good girl or boy*?
35	When she/he was 4 to 5, did she/he play any pretend or make-believe games?	When she/he was 4 to 5, did she/he play any pretend or make-believe games (e.g., playing daddy & mummy, backing a baby)?*

**a traditional African method where mothers carry babies and infants on their backs swathed in cloth*

5.4.2.2.2 Screener for Intelligence and Learning Disabilities (SCIL 14 – 17) for screening Intellectual Disability

The same participants reviewed the autism spectrum disorder and intellectual disability measures. The agreement for the SCIL 14 – 17 to give more contextual and culturally relevant examples was between 87.5% and 100%. Thus, more contextual and culturally relevant examples were given, in addition to including other relevant educational categories. There is no commercially available English version of the SCIL 14 – 17, and this study was an effort to create one. Therefore, in examining the face validity, culturally relevant words and examples were included. Question 1 on the level of education was modified to include all the different categories of educational qualifications obtained in Nigeria. One key factor was language. In

most western societies, a 'diploma' refers to a secondary school certificate, while in Nigeria, a 'diploma' refers to certificates obtained in post-secondary school. In question 3, because there are no 'services' as obtained in the West, 'services' had to be modified to exclude individuals who provided extra tutoring at home as a competitive advantage. However, where individuals visited any psychiatric facility or psychologist, these qualified as receiving service. Another example is changing the word 'GP' to 'doctor' as the term 'GP' is not utilised in Nigeria. Results of other modifications are provided in Table 14. Table 15 shows the old and modified questions for the SCIL 14 – 17.

Table 14 - List of questions and the agreed cultural examples and modifications for the SCIL 14 – 17

SCIL 14 – 17) item number	**Number of votes**	**Comments and suggested clarifications**
1	8	Type SEN in full
2	8	Add WAEC/*I*GSE/ SAT, college / monotechenic/ polytechnic/ university
3	8	Write "ID" in full; for – "service" (exclude lesson teachers)
4	8	"In case of emergency or difficult situation…"
5	7	Add Naira sign, change 6,95 to 6.50
6	7	Change GP to Doctor (can use a different context)
7	7	Change GP to Doctor (can use a different context)
8	7	Remove "say every letter"
9	7	"paper" be more specific (newspaper)
10	7	Change to "raining cats & dogs", "make hay while the sun shines", "a stitch in time saves nine"
11	7	Put in boxes
12	7	Change "deer" to "cow", use "avoid", change "hitting" to "knocking down"
13	7	Change mobile phone to "card"
14	7	Add "mins" to 15, use "detailed"

Table 15 - List of old questions and their modifications for the SCIL 14 – 17)

SCIL 14 – 17) item number	**Old question**	**Modified question**
1	Did you receive special education? Do you go to a special needs school? Did you have a SEN?	Do you receive special education? Do you go to a special needs school? Do you have a special educational need (SEN)?
2	Which school/college do you attend now?	Which school/college do you attend now, or did you attend in the past?
	None	None
	Primary school	Primary school
	Special needs school	Special needs school
	GCSE	WAEC/IGCSE/SAT
	A Level	A-level
	Polytechnic college	Polytechnic/Monotechnic/Teacher's college
	University	University
	Other	Other
3	Have you received support from a service for people with ID?	Do you receive or have you received support from a service for people with Intellectual Disability (excluding a home tutor or lesson teacher)?
4	Have you got family members or relatives who you can contact if you have a problem?	Have you got family members, relatives or friends who you can contact if you have a problem (for example a difficult situation or emergency)?
6	Imagine you are at your GP (General Practitioner) 19th of January. He wants to see you again in three	Imagine you are at your Doctors on the 19th of January. He or she wants to see you again

	weeks. When (which date) would that be?	in three weeks. When (which date) would that be?
7	Imagine you are at your GP (General Practitioner) January 3rd. He wants to see you again in three weeks. When (which date) would that be?	Imagine you are at your Doctors on the 3rd of January. He or she wants to see you again in three weeks. When (which date) would that be?
9	Do you read a paper or magazine? If so, which one?	Do you read a newspaper or magazine? If so, which one?
10	What does this mean: The apple doesn't fall far from the tree?	What does this mean: "Like father, like son?"
12	I'm going to read a few sentences for you to write in the box. Try to do this well/correct and as fast as you can. a) We are dumping the load of soil/sand at the back of our house. b) During the night the driver had to swerve/avoid hitting a deer with big antlers.	I'm going to read a few sentences for you to write in the box. Try to do this well or correct and as fast as you can. a. We are dumping a load of sand in the back garden. b. During the night the driver had to avoid knocking down a cow.
13	I'm going to ask you to read a story. Read this as quickly as you can without making mistakes. It is possible to pay for parking by text(phone). When you have parked your car, log in on your mobile using the (location) code as advertised on the signs and parking machines. When you leave you log out by phone.	I'm going to ask you to read a story. Read this as quickly as you can without making mistakes. It is possible to pay for parking with your bank card. When you have parked your car, you use your bank card to pay as advertised/displayed on the signs and parking machines. When you leave you take your receipt.

14	In this box draw a clock that says 9.45 (15 to ten). Draw this as complete/detailed as you can with hands	In this box draw a clock that says 9:45 (quarter to ten). Draw this as complete/detailed as you can with hands.

5.5 Discussion

Cross-cultural adaptation of any tool is often complicated, thus requiring careful elimination of possible construct, item, and method biases (Van de Vijver & Poortinga, 1997; Van de Vijver & Tanzer, 2004). Beyond the biases identified by Van de Vijver & Poortinga (1997) and Van de Vijver & Tanzer (2004), Peña (2007) identified another type of bias which can occur when conducting cross-cultural adaptation of screening tools called 'equivalence'. According to Peña (2007), there are four types: cultural, linguistic, metric and functional equivalence. A qualitative review of the dialogue between the nominal group participants revealed that the biases of concern were around linguistic, cultural, and functional equivalence. The linguistic equivalence ensures the consistency of words, sentences, meaning, and language used between the original and the adapted tool (Peña, 2007). One challenge with linguistic equivalence is that even when words are the same across the original and adapted tools, culture, interpretation, and word familiarity may result in potential differences in patterns of responses. In the SCIL 14 – 17, for instance, the phrase 'mobile phones' is similar in the Nigerian context; however, the function attributed to it was different. With cultural equivalence, how members of different linguistic and cultural groups interpret the underlying meaning of words or items is crucial. For instance, question 2 in the SCQ asks about 'holding to and fro conversation', which is not the norm in an average Nigerian family. With functional equivalence, both the original tool and the adapted version should allow examination of the same construct. Both versions should offer the same opportunity to demonstrate knowledge while eliciting the intended response from participants. An example of this was the observation made on question 14 of the SCQ on 'tactile' in the original version versus 'touch' in the Nigerian context. Overall, there is an interaction between the linguistic, cultural, and functional equivalence which should not be ignored in the adaptation process. Additionally, participants were concerned about the method bias (mode of

administration) and item bias, especially for the SCQ. In cultures where social interactions and dialogues are salient, dyadic administrations may be more valuable.

The modified nominal group technique was used to select the most robust screening tool for autism spectrum disorder and/or intellectual disability from those identified through a systematic review (Nwokolo et al., 2022). The cultural relevance, face validity, and content validity for use with the Nigerian adolescent were examined. The Social Communication Questionnaire is an existing measure developed in the Western environment with various translations. Three participants were familiar with the SCQ and used it often. The group reviewed the Lifetime English version with consensus reached on all the face and content validity items. On cultural relevance, the consensus was to use indigenous examples in language and activities mentioned in the SCQ. The group agreed that although the SCQ is a self-administered tool, it may be best administered as a quasi-interview questionnaire to get a more accurate response in the Nigerian context. Doing so will allow the administrator to explain potentially confusing concepts, quickly substitute examples, expound phrases, and note areas of importance or value to the respondent. This view of adapting tools to meet the specific culture and environment of intended use was captured by Soto et al. (2015).

Reviewing the Screener for Intelligence and Learning Disabilities required more depth as there currently is no English version. The group chose the SCIL 14 – 17 over the CAIDS-Q, stating that the SCIL 14 – 17 had specific questions in certain areas like mathematics and reading. The SCIL 14 – 17 was deemed more engaging and functional. Not only did they agree on the face validity, but the group also noted that the contents of the SCIL 14 – 17 tested individual abilities and the DSM-5 domains for ID (conceptual, social, and practical). The question on the level of education was expanded to include all the different curricula offered in Nigeria, including the Nigerian, British and American curricula.

To have an adapted tool that is culturally relevant, linguistically appropriate, and applicable to the environment of intended use, such as Nigeria, individuals who understand the people and are also familiar with the construct of interest need to be involved in the adaptation process. Whereas adaptation of tools includes language translation, modification of methods, clarification of concepts, and sometimes changing the content, for the tool to be genuinely relevant culturally,

the values and peculiarities of the environment of intended use should be considered (Al Maskari et al., 2018; Soto et al., 2015). For instance, the word '*ritual*' in the SCQ will elicit a different response as some people believe in idols and engage in 'rituals' (sacrificial killings) in Nigeria. Therefore, respondents will likely answer 'no' if they do not hold such beliefs or ignore the question where they feel it is a private event. The nominal group paid attention to such content and recommended that alternative wording be used or have the administrator explain the question. Similarly, on the SCIL 14 – 17, the group suggested that the type of 'service' be qualified. In Nigeria, there are no similar government-funded organisations or services like those in the West, where the tool was initially developed.

5.6 Limitations

This study focused on adapting screening tools in English, which reduces its use outside of non-English speaking African countries, and in the rural areas of Nigeria. The late arrival of one of the participants and the early exit of another meant that expert representation in those fields was not available for the entire meeting period. The small number of participants was also a limitation. It would have been ideal to have between 10 – 12 participants. Initially 2 participants per profession were invited, but the unavailability of some invitees affected this. We recommend further reliability and validity studies of the identified screening tools. While efforts to ensure the qualitative data's trustworthiness were examined, it is possible that data validity was not explored in its entirety. One possible means of exploring data validity would have been the use of investigator or theory triangulation (Guion et al., 2011). Using an evaluation team outside the researcher's group may have lent different perspectives or interpreted the data differently.

A limitation of the NGT is 'groupthink', where the more powerful or vocal individual dominates the decision; thereby leading other group members to 'agree' even in cases where the decision may be wrong. List (2001) provides guidelines to address the power disparity. Due to the stigma attached to disabilities and conversations around autism spectrum disorder and intellectual disability still at its infancy, there was no involvement of a neurodivergent person. However, a parent of a neurodivergent individual was involved. Including a neurodivergent person in future consensus studies should be considered.

5.7 Conclusion

Realising that the adaptation process beyond language translation can be complicated and challenging, using the appropriate knowledge, skill, and expertise is crucial. A group of Nigerian experts in the relevant professions were consulted to review the four identified tools for screening for intellectual disability and autism, for face validity, content validity and cultural validity, with two tools chosen for scrutiny and adaptation. Assessing some of the properties (face validity, content validity) of the screening tools using the NGT was useful. Following the recommendations and consensus of the group, the SCQ and the SCIL 14 – 17 were agreed on as measures to be validated with the Nigerian adolescents, with only a small number of adjustments to allow for different use of language, customs and environment in the Nigerian context. The SCQ and SCIL were therefore utilised in validation studies in Nigeria amongst the adolescents.

Chapter 6. Study 3a – Preliminary testing of the English Version of the Screener for Intelligence and Learning Disabilities (SCIL) [3]

6.1 Introduction

Social factors, such as parental decisions, culture, expected outcomes, and beliefs, may undermine the early assessment and diagnosis of mental health or other neurodevelopmental disorders amongst African adolescents (Dogra, 2015; Garland et al., 2004). Thus, adolescents with intellectual disabilities in African countries such as Nigeria are not screened at an early age and often go undiagnosed due to a lack of understanding of the challenges by primary caregivers and some professionals, complicated by the socio-political climate, and lack of adequate tools for screening (Franz, Chambers, von Isenburg, & de Vries, 2017; Nwokolo, Langdon, & Murphy, 2022). The United Nations Children's Fund (UNICEF) projected that 109 million persons under the age of 18 will reside in Nigeria by 2021 (UNICEF, 2014), and further predictions suggest that there will be about 1.1 billion under the age of 18 within Africa by the year 2100 (UNICEF, 2014). Neglecting their mental health and developmental needs, evidenced by the lack of copious research, poses substantial challenges (Kieling et al., 2011; Maxey & Beckert, 2017; Erskine et al., 2017). Based on scant data, the prevalence of adolescent mental illnesses (excluding intellectual disability) stands at 6.7% in Sub-Saharan Africa (Erskine et al., 2017), and the World Health Organization (WHO, 2014) mapped out strategies to address the all-round health concerns of adolescents, which included policies, collaboration between sectors, and data gathering. Given the high comorbidity of mental health disorders in this age group (Jozefiak, Kayed, Rimehaug, Wormdal, Brubakk, & Wichstrøm, 2016; Munir, 2016; Uzun Cıcek, Sarı, & Mercan Isık, 2020), screening for intellectual disability, often included within mental health subdomains, must be included within these initiatives.

[3] A version of this paper was submitted to the Journal of Intellectual Disability Research, in March 2023, as Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E., (submitted). Preliminary testing of the English Version of the Screener for Intelligence and Learning Disabilities (SCIL) amongst adolescents in Nigeria.

A diagnosis of intellectual disability must meet the diagnostic criteria within three domains: a) intellectual functioning, b) adaptive skills, and c) onset in the developmental period, according to the International Classification of Diseases (ICD-11; World Health Organization [WHO], 2020). The administration of complete intelligence tests is usually lengthy, time-consuming, costly and requires trained professionals. Consequently, screening tools were developed and used to save time and cost. Screening helps with the identification of persons who *may* have an intellectual disability and need further assessment. Early identification of adolescents suspected to have an intellectual disability is crucial for adequate classification, diagnosis, and tailoring of the right support and environment for them to thrive (Franz et al., 2017; Matson, Rieske, & Tureck 2011). In Africa, in countries such as Nigeria, a lack of adequate and validated screening tools and a low level of awareness among parents and professionals have been identified as barriers to assessment (Franz et al., 2017; Nwokolo et al., 2022).

To identify available time and cost-efficient screening tools for use with Nigerian adolescents, Nwokolo et al. (2022) conducted a systematic review (Chapter 4) of such tools examining their cultural appropriateness and psychometric properties. A total of six tools were identified: (1) Hayes Ability Screening Index (HASI), (2) the Learning Disability Screening Questionnaire (LDSQ), (3) the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q), (4) the Slosson Intelligence Test (SIT), (5) the Screener for Intelligence and Learning Disabilities (SCIL), and (6) the Quick Test (QT). After assessing the evidence on the six tools, two of them (the SCIL and the CAIDS-Q) were selected for further review by a focus group (Chapter 5) of Nigerian experts (Nwokolo et al., in press). The group of experts examined the face and content validity of both tools. After the focus group consultations in Nigeria, the adolescent version of the Screener for Intelligence and Learning Disabilities (SCIL 14-17) was selected. The group agreed the SCIL had more practical and functional items for screening intellectual disabilities than the CAIDS-Q. The SCIL is a short, 14-item tool for identifying individuals with a level of general intellectual functioning that falls within and below the "very low" range (Nijman et al., 2018). The adolescent version of the SCIL was written in Dutch and translated into English, and minor changes were made following the Nigerian focus group's recommendations regarding the cultural appropriateness of examples in the SCIL for Nigerian adolescents. Lower scores on the SCIL indicate suspicion of possible intellectual disability, with

further assessments and diagnosis recommended. Administration of the SCIL is time efficient and does not require any costly or special training.

To ascertain the usefulness of the English version of the adolescent SCIL in Nigeria, as translated and adapted by Nwokolo et al. (in press), 209 adolescents and young adults were invited to complete the SCIL and undergo measures of their adaptive behaviour and level of general intellectual functioning. This study aimed to a) examine the component structure of the SCIL and reduce dimensions as required, b) examine the internal consistency, discriminative, and convergent validity of the SCIL, c) derive an appropriate cut-off score based upon sensitivity and specificity and d) derive the positive and negative predictive values.

Given these aims, the recruitment of participants from the relevant centres was purposive to allow the inclusion of those with and without an intellectual disability.

6.1.1 Skill and Experience of the Research Assistants

This study involved 3 research assistants. One was a clinical psychologist with over eight years of experience as a child and adolescent psychologist who worked at the Child and Adolescent Mental Health Hospital. The second was a developmental psychologist, an independent practitioner who had a Master's with two years of experience working with children and adolescents. The clinical psychologist was referred to the main researcher by the Head of Department, Psychology at the Federal Psychiatric Hospital in Lagos, Nigeria, while the developmental psychologist was contacted through the main researcher's contacts. Both psychologists administered psychological tests as part of their routine work and administered the Wechsler Intelligence Scale for Children Fifth Edition (WISC-V) and the Wechsler Adult Intelligence Scale Third Edition (WAIS-3). The third research assistant was a Qualified Autism Service Practitioner- Supervisor (QASP-S). The QASP-S is a certificate obtained from the Qualified Applied Behavior Analysis Credentialing Board (QABA), a US-based organisation recognised in Nigeria. The credential holder has a specialised training in applied behaviour analysis (ABA), has demonstrated clinical skills to utilise ABA across a range of settings and with a variety of clients and has advanced knowledge of autism and experience in working with autistic individuals in various settings. Study details were shared with the research assistants. The lead researcher provided one full day of training to the QASP-S on administering the

consent forms, information sheets and screening tools, followed by three days of demonstration. Following this, the QASP-S was observed while implementing the procedure. No procedural errors were observed.

6.2 Methods

6.2.1 Design

A between-groups design was used with two groups of participants: adolescents and young people thought to have intellectual disabilities and those thought not to have such disabilities.

6.2.2 Participants

The study took place within three geopolitical zones in Nigeria, namely, Enugu, Abuja, and Lagos. In the 'suspected intellectual disability group', an adolescent or young person was eligible to participate in this study if they a) were between 11- and 26 years old, b) were identified by a medical doctor as possibly having an intellectual disability, and/or c) attended a special school or a special day centre. Other participants of the same age, who were thought unlikely to have an intellectual disability, were also recruited. Participants were recruited from day centres, special schools, child and adolescent mental health care services, local community organisations, places of worship and public advertisement. Initially, 245 adolescents were invited to take part in this study, and 35 declined to participate or did not respond to further attempts to contact them; finally, 210 adolescents (M_{age} = 15.88 years; Mdn_{age} = 15.29 years; SD = 3.69; Min: 10.90 years; Max: 26.96 years; 41% female and 59% male) took part in this study. The age distribution was categorised as follows: 11 – 13-year-olds (n = 76; 36.2%), 14 – 15-year-olds (n = 42; 20%), 16 – 17-year-olds (n = 51; 24.3%) and 18 years and above (n = 41; 19.5%).

6.2.3 Measures

6.2.3.1 Screener for Intelligence and Learning Disabilities (SCIL)

The SCIL (Nijman et al., 2018; Geijsen et al., 2018) is a standardised 14-item questionnaire used to screen intellectual disabilities and takes 10 – 15 minutes to complete. The measure was developed and used for adolescents in the Netherlands, and it was translated and back translated into English and subject to minor modifications to ensure it was culturally suitable for Nigerians

(described in Chapter 5). A further description of the SCIL items is available elsewhere (Geijsen et al., 2018).

6.2.3.2 Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V)

The WISC-V (Wechsler, 2014) is an individually administered intelligence test for children aged 6 to 16. Administration of the WISC-V takes between 45 to 65 minutes. The child's Full-Scale Intelligence Quotient (IQ) is generated from 7 of the primary subtests – Verbal Comprehension Index (2 subtests), Visual Spatial Index (1 subtest), Fluid Reasoning Index (2 subtests), Working memory Index (1 subtest) and Processing Speed Index (1 subtest).

6.2.3.3 Wechsler Adult Intelligence Scale, Third Edition (WAIS-III)

The WAIS-III (Wechsler, 1997) is standardised for use with older adolescents and adults and was used to calculate Full-Scale IQ for participants who were older than 16 years. Both Verbal and Performance IQ were also calculated. Administration can take around 60 minutes.

6.2.3.4 Vineland Adaptive Behavior Scales, Third Edition (VABS-3)

The VABS-3 is a standardised semi-structured interview to index adaptive behaviour. It can be completed with an individual, their carer/parent, or a teacher. The carer/parent domain level form was used within this study as it has been recommended for research purposes (Pepperdine & McCrimmon, 2018; Sparrow, Cicchetti, & Saulnier, 2016).

6.2.4 Procedure

A favourable ethical opinion was obtained from the University of Kent, Tizard Centre Ethics Committee, the National Health Research Ethics Committee (NHREC; NHREC/01/01/2007-16/09/2019) and the Federal Neuro-Psychiatric Hospital, Yaba, Lagos, Nigeria (FNPH/HREC/20/09; Appendices 1 to 4). All participants were provided with written information about participating in this study, assent, and consent forms, including easier-to-read versions (Appendix 7, 8, 9, 11, 16, 17 & 18). Parental informed consent was sought for those aged under 18 years of age (Appendix 10). All participants were given the feedback form which contained information of whom to contact should they have any concerns about the research such as withdrawing (Appendix 19). Participants were encouraged to take breaks as needed during testing.

Each participant was invited to complete the SCIL and either the Wechsler Adult Intelligence Scale, 3rd Edition (WAIS-III) or the Wechsler Intelligence Scale for Children, 5th Edition (WISC-V), depending on age. The Vineland Adaptive Behavior Scales, 3rd Edition (VABS-3) were completed with the parent or caregiver with sufficient developmental knowledge about the participant. The WAIS-III and the WISC-V were administered by a psychologist, while the VABS-3 and SCIL were administered by a qualified autism service practitioner supervisor trained on the assessments.

6.2.5 Data Analysis

6.2.5.1 Overview

The Statistical Package for the Social Sciences – IBM SPSS version 26 was used for the analysis. Principal component analysis (PCA) was used to examine the component structure of the English version of the SCIL to determine whether any items should be removed, noting that this was conducted previously using the Dutch version (Geijsen et al., 2018). PCA was performed with a Promax rotation with Kaiser normalisation because a correlation between the components was expected. First, individual items were retained if they correlated at least .30 with another item. Second, the remaining items were then retained if item communalities were at least .30 (Floyd & Widaman, 1995) and the criterion of at least .60 for the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was met. Components with eigenvalue >1 were retained.

Following this, discriminative and convergent validity and internal consistency were examined. Convergent validity was assessed by examining the extent to which the scores on the SCIL correlated with Full-Scale IQ and the VABS-3. Receiver Operating Characteristic (ROC) analyses were used to calculate the Area under the Curve (AUC) to examine how well the SCIL identified participants with and without an intellectual disability, with reference to sensitivity and specificity to identify optimal cut-offs to examine discriminative validity. Internal consistency was considered by calculating Cronbach's α. The positive predictive value (PPV) and the negative predictive value (NPV) were determined using the cross-validation of the classification results.

6.2.5.2 IQ and VABS-3 cut-offs for Intellectual Disabilities

Participants were classified as having an intellectual disability using ICD-11 (WHO, 2020) diagnostic criteria for disorders of intellectual development. Specifically, those with a level of general intellectual functioning and adaptive behaviour that fell at least two standard deviations or more below the mean were classified as having an intellectual disability. However, there is evidence (Wicherts, Dolan, & van der Maas, 2010; Ani & Grantham-McGregor, 1998; Nenty & Dinero, 1981; Ashem & Janes, 1978, Nwanze & Okeowo, 1980 and Fahrmeier, 1975) to show that the average IQ of Africans on Western IQ tests is at least one standard deviation lower than Western normative data. Questions and discussions around the validity and cultural appropriateness of IQ tests such as the WISC and WAIS developed in the West (predominantly with White populations) and used in Africa or with the Black populations continue to occur (Kamin, 2006; Rushton & Jensen; 2006; Bakare, Ubochi, Okoroikpa, Aguocha & Ebigbo, 2009; Shuttleworth-Edwards, Donnelly, Reid & Radloff, 2004; Shuttleworth-Edwards, Kemp, Rust, Muirhead, Hartman & Radloff, 2004). Recognising that environment, culture, exposure to Western cultures, level and quality of education, and various other factors contribute to an estimate of the level of general intellectual functioning (Bakare et al., 2009; Shuttleworth-Edwards, Kemp, Rust et al., 2004), it would be "illogical to use scores on such tests to infer the level of innate ability possessed by people in non-Western cultures or as a basis for making judgmental statements of a superior-subordinate nature about their performance" (Bakare, 1972, p.362). However, Wicherts et al. (2010) systematically reviewed published empirical data on various Western IQ tests with Western norms regarding the performance of African populations. The tests included the Kaufman-Assessment Battery for Children (K-ABC), the Wechsler scales (WAIS & WISC), the Draw-A-Man (DAM) test, and several others, with the goal of estimating the average IQ of samples of normal and healthy Africans. Their results showed that the average IQ of Africans was approximately 82 when compared to UK norms.

Regarding the Vineland Adaptive Behavior Scales (VABS), several studies have compared the VABS to measures of adaptive behaviour that were developed in African countries (du Toit, Van der Linde & Swanepoel, 2021a; du Toit, Van der Linde & Swanepoel, 2021b; de Beer, Krüger, Van der Linde, Eccles & Graham, 2020; Allen et al., 2014). Unfortunately, the authors of these studies did not report the data captured using the VABS to allow for conclusions about how well

this instrument performs when used within Africa (Beer et al. 2020; du Toit et a. 2021a, 2021b; Allen et al. 2014). Douglas (2017) concluded that the VABS can identify persons with intellectual disability under 22-years of age amongst sexual abuse victims in South Africa but did not report their actual VABS data making it difficult to examine how well the VABS performed.

Considering the issues mentioned above and our work's cultural sensitivity, we made use of two different IQ cut-offs for identifying participants with an intellectual disability and completed our analysis twice. Initially, we used the established, **unadjusted**, Western criterion of FSIQ <70 (i.e., 2 SD below 100); then, secondly, the **FSIQ cut-off was adjusted** based on the work of Wicherts et al. (2010) to FSIQ score <52 (i.e., 2 SD below 82). There is a lack evidence about the performance of the VABS-3 when used to identify intellectual disability within Africa. Considering this, we opted to retain and use the established cut-off of <70 (i.e., 2 SD below 100) for the VABS for deciding whether a participant had an intellectual disability for our analysis using both the adjusted and unadjusted Full Scale IQ cut-off. In a small number of cases (n = 8), the VABS-3 Composite score was above 70 while the FSIQ was below 70: in these cases, they were allocated to the 'no intellectual disability, unadjusted cut-off' group. Also, in a further small number of cases (n = 9), the FSIQ was above 52 while the VABS-3 Composite score was below 70: in these cases, they were allocated to 'no intellectual disability, adjusted cut-off' group.

6.2.5.3 Missing Data

Thirty-seven participants (30 with WISC-V, 7 with WAIS-III) could not complete the assessment of their level of general intellectual functioning, as they were scoring **below the floor of the test**, possibly due to the degree of intellectual disability, while one of these participants was also unable to score on the SCIL and the VABS-3. Therefore, thirty-seven participants were excluded from analyses using the Weschler scales, while 1 participant was excluded from the analysis using VABS Composite scores and SCIL analyses.

6.3 Results

6.3.1 Descriptive Statistics

SCIL scores for all participants ranged from 0 to 28 points, M = 14.01; Mdn = 16.00; SD = 9.4. SCIL scores did not differ between the sexes, $t_{(207)} = -1.818$, p = .07. Descriptive statistics for the SCIL by age group for the whole sample are in Table 16 and the descriptive statistics for the VABS-3 are in Table 17. There was no significant correlation between age and the SCIL scores $r_{(209)} = .13$, p = .06. Likewise, Tables 17 and 18 show the mean and SD for VABS-3 Composite scores for the whole group and for FSIQ scores for the whole group. There was significant correlation between the VABS-3 Composite scores and the FSIQ scores $\mathbf{r}_{(173)} = .74$, $p < .001$ (excluding the 37 who scored below the floor on the Weschler Scales).

When the initial criterion (FSIQ <70) for very low intellectual functioning was applied to the whole sample, 68.57% (N = 144) were classified as having an intellectual disability, 30.95% (N = 65) as not having an intellectual disability and 0.5% (N = 1) missing. However, applying the criterion of FSIQ <52 to the whole sample, 70% (N = 147) of participants were reclassified as not having an intellectual disability and 29.52% (N = 62) with an intellectual disability. The VABS-3 Composite scores criterion of <70 classified 69.05% (N = 145) as not having an intellectual disability and 30.48% (N = 64) with an intellectual disability, and 0.5% (N = 1) missing.

It is important to note these figures do NOT give the prevalence of intellectual disability in Nigeria since the two groups were specifically chosen so as to represent likely intellectual disability and unlikely intellectual disability. As previously mentioned, 37 participants did not complete the FSIQ assessment due to their level of intellectual functioning and were subsequently excluded from the analyses.

Table 16 - Descriptive statistics of SCIL scores by age groups

Age Groups	n	Mean	SD	Range	Min	Max
11 to 13 years old	76	12.22	9.29	27	0	27
14 to 15 years old	42	15.07	8.94	26	0	26
16 to 17 years old	51	14.43	9.28	28	0	28

18 years old & above	40	15.75	9.98	28	0	28
Total score	209	14.01	9.40	28	0	28

Table 17 - Descriptive statistics of VABS Composite scores <70 by age groups

Age Groups	n	Mean	SD	Range	Min	Max
11 to 13 years old	76	81.57	25.71	104	20	124
14 to 15 years old	42	78.98	24.03	96	25	121
16 to 17 years old	51	77.41	21.96	80	32	112
18 years old & above	40	77.18	25.32	76	37	113
Total	209	79.19	24.33	104	20	124

Table 18 - Descriptive statistics of FSIQ scores by age groups (37 participants did not complete the FSIQ assessment as they were scoring below the floor of the test)

Age Groups	n	Mean	SD	Range	Min	Max
11 to 13 years old	57	79.04	20.05	85	42	127
14 to 15 years old	37	72.97	23,51	89	40	129
16 to 17 years old	43	74.44	18.56	65	40	105
18 years old & above	36	77.19	25.07	93	45	138
Total	173	76.21	21.551	98	40	138

Thirty-seven percent of participants were initially categorised as having an intellectual disability as they had a Full-Scale IQ and VABS Composite Score of <70 (Table 19). Adjusting the Full-Scale IQ cut-off to <52 resulted in 16% of our participants being categorised as having an intellectual disability (Table 20).

Table 19 - Descriptive statistics of the WISC-V and WAIS-III scores for participants in the two groups (with & without intellectual disabilities) using an unadjusted criterion for identification of intellectual disability FSIQ score <70 (i.e., 2 SDs below 100) excluding the 37 who scored below the floor of the test and the VABS-3 Composite score <70 for the whole group.

	With Intellectual Disabilities					Without Intellectual Disabilities				
	n	Min	Max	Mean	SD	n	Min	Max	Mean	SD
Unadjusted WISC-V FSIQ	41	40	69	53.51	9.88	80	70	129	88.31	14.18
Unadjusted WAIS-III FSIQ	23	45	69	53.70	8.01	29	71	138	92.79	14.90
Total Participants	64	40	69	53.58	9.19	109	70	138	89.51	14.44
VABS Composite Score	64	20	70	49.55	13.42	145	58	124	92.28	14.61

Finally, the participants were divided again, this time using the adjusted cut-off of IQ 52, so as to give the mean scores and SD for the 'Intellectual Disabilities, adjusted cut-off' and 'no Intellectual Disabilities, adjusted cut-off' groups, as described above in the Method section. Table 20 gives the resulting FSIQ scores for these two groups.

Table 20 - Descriptive statistics of the WISC-V and WAIS-III scores for participants in the two groups (with & without intellectual disabilities) using an unadjusted criterion for identification of intellectual disability FSIQ score <70 (i.e., 2 SDs below 100) excluding the 37 who scored below the floor of the test.

	With Intellectual Disabilities					Without Intellectual Disabilities				
	N	Min	Max	Mean	SD	N	Min	Max	Mean	SD
Adjusted WISC-V FSIQ	18	40	51	43.90	3.97	103	52	129	82.23	17.10
Adjusted WAIS-III FSIQ	9	45	48	45.78	1.09	43	52	138	81.72	20.47
Total Participants	27	40	51	45.52	3.39	146	52	138	82.08	18.09

6.3.2 Principal Component Analysis (PCA)

The correlation matrix revealed that all items of the SCIL correlated at $r > .30$ with more than one other item; thus, all 14 items of the SCIL were initially retained. Multicollinearity criteria ($r > .8$) analysis showed that all correlations were less than .8. The KMO was .94, and all KMOs

for individual items were greater than .87, which is well above the acceptable limit of .5 (Field, 2013; Kaiser, 1974). Bartlett's test of sphericity was significant, $\chi 2(91) = 1947.49$, $p < .001$, which implies that the items are correlated. Therefore, all items were retained. The scree plot was slightly ambiguous, showing inflexions that would justify retaining three factors. However, two factors had eigenvalues >1 and explained 63.56% of the variance. The Pattern Matrix and Structure Matrix are in Table 21. A full description of the SCIL items is in Appendix 23. The 9 items (1, 2, 3, 4, 8, 9, 10, 12, 13) on component 1 were judged to relate to Education, Social Contacts, and Comprehension (EsoC), while the 5 items (5, 6, 7, 11, 14) on component 2 were judged to relate to Arithmetic and Numbers (ArN).

Table 21 - Pattern and Structure Matrix of the SCIL

	Pattern Matrix		Structure Matrix	
	1	2	1	2
Item 1	**0.79**	0.05	**0.83**	0.60
Item 2	**0.83**	0.02	**0.84**	0.59
Item 3	**0.60**	0.24	**0.77**	0.66
Item 4	**0.93**	-0.40	**0.65**	0.25
Item 5	-0.06	**0.68**	0.41	**0.64**
Item 6	-0.27	**0.97**	0.41	**0.78**
Item 7	-0.11	**0.92**	0.54	**0.85**
Item 8	**0.92**	-0.11	**0.84**	0.53
Item 9	**0.38**	0.24	**0.55**	0.50
Item 10	**0.59**	0.31	**0.81**	0.72
Item 11	0.41	**0.49**	0.75	**0.77**
Item 12	**0.47**	0.46	**0.79**	0.79
Item 13	**0.72**	0.25	**0.90**	0.75
Item 14	0.30	**0.53**	0.67	**0.74**

Note: Promax Rotation with Kaiser normalisation. Component 1: Eigenvalue 7.77 (% of variance = 55.52), $\alpha = .92$. Component 2: Eigenvalue 1.13 (% of variance = 8.04), $\alpha = .84$.

6.3.3 Internal Consistency

The overall Cronbach's alpha value of .94 indicated the high internal consistency of the SCIL items. Component 1 had a value of .92, while component 2 was .84, which were both high.

6.3.4 Convergent Validity

There was a significant positive correlation between total SCIL score and level of general intellectual functioning (judged by the FSIQ score), $r_{(173)} = .81$, $p < 0.001$, indicating a large effect size (Cohen, 1992). Correlation between the total SCIL score and Full-Scale IQ estimated using the WAIS-III, $r_{(52)} = .86$, $p < 0.001$, and the WISC-V, $r_{(121)} = .79$, $p < 0.001$, was high. Similarly, there was a significant positive correlation between the total SCIL score and VABS-3 ABC Composite Score, $r_{(209)} = .84$, $p < 0.001$.

6.3.5 Sensitivity and Specificity *(FSIQ and VABS-3 Composite Scores <70)*

At the suggested cut-off score of 15 (Nijman et al., 2018), the AUC was .81, $p < 0.001$, 95% CI [.75, .86], sensitivity = 1 and specificity = .55, applicable to the entire sample. However, exploring lower cut-off scores of 10, 11 and 12 did not improve the specificity values (.44, .52, and .54, respectively), although sensitivity remained at 1. For this study, there was no difference in categorising participants with likely or unlikely intellectual disability using the cut-off scores of 13 and 14 versus the suggested 15. As such, the sensitivity and specificity using 13 or 14 were not explored. Using 15 as the cut-off, the AUC across the age groups was moderate: 11 – 13-year-olds, AUC = .85, $p < 0.001$, 95% CI [.76, .93], $N = 76$, 14 – 15-year-olds, AUC = .74, $p < 0.001$, 95% CI [.60, .89], $N = 42$, 16 – 17-year-olds, AUC = .79, $p < 0.001$, 95% CI [.67, .91], $N = 51$, 18 years and above, AUC =.83, $p < 0.001$, 95% CI [.71, .96], $N = 40$.

6.3.6 Positive Predictive Value (PPV) and Negative Predictive Value (NPV) *(FSIQ and VABS-3 Composite Scores <70)*

At the suggested SCIL cut-off score of 15, (Nijman et al., 2018), the PPV was .62 and the NPV was 1.

6.3.7 Sensitivity and Specificity *(FSIQ <52 and VABS-3 Composite Scores <70)*

When the adjusted FSIQ (i.e., with the cut-off of 52) and VABS-3 Composite scores (i.e., 2 SDs below 100) were used, using the suggested SCIL cut-off score of 15 (Nijman et al., 2018),

sensitivity = 1 and specificity = .78 for the entire sample. Both measurement properties met the minimum standard (Glascoe, 2005). The AUC was .89, $p < 0.001$, 95% CI [.85, .93]. Also, using 15 as the cut-off, the AUC across the age groups was large: 11 – 13-year-olds, AUC = .83, $p < 0.001$, 95% CI [.75, .92], $N = 76$, 14 – 15-year-olds, AUC = .95, $p < 0.001$, 95% CI [.88, 1.02], $N = 42$, 16 – 17-year-olds, AUC = .90, $p < 0.001$, 95% CI [.82, .98], $N = 51$, 18 years and above, AUC = .91, $p < 0.001$, 95% CI [.83, 1.00], $N = 40$. To determine the best cut-off score, rather than selecting an arbitrary figure, lower cut-offs were explored by stepwise reduction. Lowering the cut-off score to 10, 11 and 12 improved the values; however, a cut-off score of 10 gave the best result. A cut-off score of 13 did not yield different results from using 15 for this study and was not explored further. The AUC indicates the discriminative ability of the SCIL; a perfect tool would have an AUC of 1, and the AUCs in this study ranged from .83 to .95, indicating good discriminative validity. The AUC, PPV, NPV, sensitivity and specificity associated with each cut-off score are shown in Table 22 – figures in bold indicate measurement properties at the applicable cut-off scores.

6.3.8 Positive Predictive Value (PPV) and Negative Predictive Value (NPV) *(FSIQ <52 and VABS-3 Composite Scores <70)*

Using the cut-off of 10, the SCIL had a PPV = .66 and NPV = 1, and with a cut-off of 11, PPV = .65 and NPV = 1, while at a cut-off of 12, PPV = .61 and NPV = 1. These show that the SCIL can correctly identify those with and without intellectual disabilities. However, the most suitable cut-off was determined to be 10.

Table 22 - Sensitivity, Specificity, PPV and NPV for the various potential cut-off scores of the SCIL using the adjusted mean FSIQ Scores

	SCIL Cut-off Score <10					SCIL Cut-off Score <11					SCIL Cut-off Score <12				
	Total (n=209)	11 - 13 years (n = 76)	14 - 15 years (n = 42)	16 - 17 years (n = 51)	18 years & above (n = 40)	Total (n=209)	11 - 13 years (n = 76)	14 - 15 years (n = 42)	16 - 17 years (n = 51)	18 years & above (n = 40)	Total (n=209)	11 - 13 years (n = 76)	14 - 15 years (n = 42)	16 - 17 years (n = 51)	18 years & above (n = 40)
PPV	**0.66**	**0.53**	**0.92**	0.65	0.79	0.65	0.51	0.86	**0.65**	**0.79**	0.61	0.51	0.75	0.75	0.69
NPV	**1.00**	**1.00**	**1.00**	1.00	1.00	1.00	1.00	1.00	**1.00**	**1.00**	1.00	1.00	1.00	1.00	1.00
Sensitivity	**1.00**	**1.00**	**1.00**	1.00	1.00	1.00	1.00	1.00	**1.00**	**1.00**	1.00	1.00	1.00	1.00	1.00
Specificity	**0.83**	**0.70**	**0.97**	0.85	0.90	0.81	0.68	0.93	**0.85**	**0.90**	0.78	0.68	0.87	0.83	0.83
AUC ($p < 0.001$, 95% CI)	0.91	0.85	0.98	0.93	0.95	0.91	0.85	0.97	0.93	0.95	0.89	0.84	0.93	0.91	0.91
Lower	0.88	0.77	0.95	0.85	0.88	0.87	0.76	0.91	0.85	0.88	0.85	0.76	0.86	0.84	0.83
Upper	0.95	0.93	1.02	1.00	1.02	0.95	0.93	1.02	1.00	1.02	0.93	0.93	1.01	0.99	1.01

Figures in bold indicate applicable properties at each cut-off score for the entire population and per age group.

6.4 Discussion

The aims of this study were to examine a) the component structure of the SCIL, b) the internal consistency, discriminative, and convergent validity of the SCIL, c) the likely appropriate cut-off score based upon sensitivity and specificity, and d) the positive and negative predictive values. Early identification of intellectual disabilities is essential for significant progress in intervention, educational support, and policy (Luckasson & Schalock, 2013; Schalock & Luckasson, 2013), but many adolescents in Nigeria have gone undiagnosed in early childhood (Franz et al., 2017).

6.4.1 Component Structure

Following an examination of the component structure of the SCIL, all 14 items were retained. Two components were derived, which we labelled (1) Education, Social Contacts, and Comprehension, and (2) Arithmetic and Numbers. Geijsen et al. (2018) previously used PCA to examine the component structure of the SCIL and retained 4 components. This disparity in findings may be due to the sampling environment, sample characteristics and the use of a different version of the SCIL. Geijsen et al. (2018) study was conducted with Dutch participants aged between 18 and 63 years in police custody, whereas participants for this study were Nigerian adolescents aged between 11 and 26 years from schools, CAMHS and the public. Also, this study employed the adolescent SCIL, whereas Geijsen et al. (2018) used the adult version of the SCIL.

6.4.2 Internal Consistency and Convergent Validity

The result indicated that the internal consistency of the SCIL was excellent. Convergent validity measures the relationship between two related constructs; in this study, it was the relationship between the SCIL scores, FSIQ and VABS-3 Composite scores. For this, positive relationships were found between FSIQ, the Composite scores and SCIL scores.

6.4.3 Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value (FSIQ and VABS-3 Composite scores <70)

ROC analyses were conducted separately for the entire sample and for each age group. The specificity did not meet the minimum criteria of 70% at the suggested cut-off of 15 (Nijman et al., 2018), while sensitivity met the criteria (Glascoe, 2005). At the suggested cut-off of 15,

about 55% of participants would have been excluded as not having a possible intellectual disability. Lowering the cut-off to 10, 11 and 12 did not improve the results. The resultant low specificity at FSIQ <70 and VABS-3 Composite score <70 for the SCIL is possibly due to the comparison of the mean IQ of our study participants to the Wechsler scales based on UK norms. These cut-off scores are not likely to represent deficits in cognitive abilities, given that the FSIQ and VABS-3 Composite scores used were based on the UK norm. Psychometric assessments in cross-cultural environments are problematic, considering the role that culture, beliefs, language, ethnicity, exposure, and quality of education play in test-taking.

6.4.4 Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value (FSIQ <52 and VABS-3 Composite Scores <70)

When using adjusted norms (FSIQ <52 and VABS-3 Composite <70), ROC analyses were conducted separately for the entire sample and for each age group, and in all cases, the sensitivity, specificity, and NPV met the minimum acceptable criteria of 70% accuracy for screening tools (Glascoe, 2005) while the PPV was within an acceptable range. An entire sample cut-off score of 10 was deemed appropriate for this study. Specific cut-offs were derived for the different age groups: 11 – 13 years (10), 14 – 15 years (10), 16 – 17 years (11), and 18 years and above (11). The sensitivity and specificity for all met the minimum criteria (70). In assigning these cut-offs, consideration was given to the distinctions between sensitivity and PPV and between specificity and NPV in a screening and clinical context (Trevethan, 2017; Akobeng, 2007). Classifying participants based solely on sensitivity and specificity values differ from classifying them in combination with the PPV and NPV. PPV and NPV are influenced by the condition's prevalence and depend on the population being investigated. Thus, in selecting 10 as the cut-off score, a combination of the PPV, NPV, sensitivity and specificity was used. Nijman et al. (2018) previously expressed concerns about the appropriateness of using the SCIL with the younger population (12 to 13-year-olds). However, our findings indicated that the SCIL could identify younger participants with and without intellectual disabilities. It is important to note that people attain developmental milestones at different times; having different cut-offs for the different age ranges is useful, noting that the cut-off score associated with each age group was near the overall cut-off of 10.

6.5 Limitations

There are limitations to this study. First, the SCIL was used with a sample of adolescents drawn mainly from large urban cities in Nigeria, which may have introduced some bias, as this sample is not representative of the whole population of Nigeria. Secondly, the sample size was reduced by 37 due to some individual's FSIQ scores being below the floor of the test. However, our sample size remained appropriately large for our chosen analyses. Thirdly, the test-retest reliability of the SCIL was not examined. Fourthly, the VABS-3, has not been validated for use in Africa or Nigeria, thus, there is uncertainty about the cut-off score that should be used to identify those with intellectual disabilities.

The degree of uncertainty regarding the cultural sensitivity of the VABS-3 is currently unknown; however, in the absence of an alternative measure for adaptive behaviour and considering the observations from Tan, Reich, Hart, Thuma, & Grigorenko (2014) and other VABS studies previously mentioned, we utilised the VABS-3. Further VABS-3 validation work in Nigeria and Africa is needed. Lastly, the adjustments to the mean FSIQ used to identify participants with intellectual disabilities may be potentially problematic and sensitive in nature, which we wish to recognise. However, as previously mentioned, there are studies to support the adjustments. Despite these limitations, our findings indicated that the SCIL is a valid screening tool for intellectual disability in Nigerian adolescents.

As previously mentioned, there are limitations to the use of sensitivity and specificity, which can be overcome by also examining the PPV and NPV. Trevethan (2017, p.4) considered that "… sensitivity and specificity indicate the concordance of a test with respect to a chosen referent, while PPV and NPV, respectively, indicate the likelihood that a test can successfully identify whether people do or do not have a target condition, based on their test results." Our results demonstrated that the English translation of the SCIL can be used for screening for intellectual disabilities in the Nigerian adolescent population. The reservations regarding its usefulness in younger adolescents (below 14 years) raised by Nijman et al. (2018) can be overcome by examining the PPV and NPV in addition to the sensitivity and specificity for the specific age groups. For this study, sensitivity, specificity, PPV and NPV were adequate (1, 0.70, 0.53 & 1 respectively) for those below 14 years.

6.6 Conclusions

The findings indicated that the English SCIL has good construct, convergent and discriminative validity. The SCIL can offer a valuable means of identifying adolescents likely to have an intellectual disability to facilitate intervention at an earlier stage (Franz et al., 2017), provide targeted support (Kieling et al., 2011), and help ensure referrals for further diagnosis. As a simple and quick screening tool, further research utilising the English version of the adolescent SCIL in more Nigerian cities and in other English-speaking African populations is recommended.

Chapter 7. Study 3b – Validation of the Social Communication Questionnaire (SCQ) [4]

7.1 Introduction

The diagnosis of autism spectrum disorder (ASD), a condition characterised by restricted and repetitive behaviours and social and communication deficits, has become increasingly common (Wing & Potter, 2002). With no known cure, a series of studies have found that lifetime costs for individuals with ASD can run into hundreds of thousands of dollars (Horlin, Falkmer, Parsons, Albrecht & Falkmer, 2014; Penner, Rayar, Bashir, Roberts, Hancock-Howard & Coyte, 2015; Sampaio, Feldman, Lavelle & Skokauskas, 2021; Rosenberg, Landa, Law, Stuart & Law 2011). However, early diagnosis and intervention have been shown to produce progress in independent functioning levels, development rate and access to effective services (James & Smith, 2020; Delehanty, Lee, Hooker, Cortese & Woods, 2020; Nadel & Poss, 2007). Nevertheless, appropriate and prompt diagnoses are crucial for accessing such intervention services early in life to capitalise on these gains.

Screening and diagnosis of ASD are feasible in very young children and are recommended as best practice; however, this has not been the norm in Nigeria, as most individuals with ASD are not diagnosed until after five years of age, and many are never diagnosed (Franz, Chambers, von Isenburg & de Vries, 2017; Bello-Mojeed, Bakare & Munir, 2013; Bello-Mojeed, Omigbodun, Bakare & Adewuya, 2017). Different factors, such as low level of awareness, limited availability of qualified professionals, cultural differences, and access to standardised tools, affect the early assessment and diagnosis of developmental disorders such as autism spectrum disorder (ASD) amongst African adolescents (Franz et al., 2017; Bello-Mojeed et al., 2013; Bello-Mojeed et al., 2017; Burkett, Morris, Manning-Courtney, Anthony & Shambley-Ebron, 2015). Therefore, adolescents with ASD in African countries such as Nigeria often go undiagnosed.

[4] A version of this paper was submitted to the journal Autism Research, in March 2023, as Nwokolo, E. U., Murphy, G. H., & Langdon, P. E. Validation of the Social Communication Questionnaire (SCQ) amongst Nigerian adolescents.

Currently, ASD diagnoses in Nigeria do not include the use of a standardised schedule (Oshodi et al., 2017; Bakare et al., 2022). Whereas a clinical assessment of autism spectrum disorder can be given based on history taking, observation and use of the DSM-5 criteria by healthcare professionals, a confirmatory diagnosis using an acceptable gold standard schedule is required for better certainty (Zeidan, Fombonne, Scorah, Ibrahim, Durkin, Saxena, et al., 2022; McCarty & Frye, 2020). The cost of acquiring such a tool is not only prohibitive, but the administration also requires trained professionals, both of which are scarce in Nigeria (Abubakar, Ssewanyana, & Newton, 2016). Therefore, level two screening tools can be used to save time and cost, and such screening would help with the immediate identification of individuals at risk of ASD. In the African context, in countries such as Nigeria, however, limited availability of age-appropriate screening and validated screening tools, as well as low levels of awareness among parents and professionals, have been identified as barriers to assessment (Franz et al., 2017; Nwokolo, Langdon & Murphy, 2022).

A systematic review as described in Chapter 4, was conducted to identify available brief and cost-efficient screening tools for use with Nigerian adolescents (Nwokolo et al., 2022), aiming to judge their cultural appropriateness and assess the psychometric properties of available tools. A total of 12 screening tools for ASD were identified through this review. The tools were Social Communication Questionnaire (SCQ), Childhood Autism Rating Scale (CARS), Child Behavior Checklist (CBCL), Pervasive Developmental Disorder in Mentally Retarded Persons Scale (PDD-MRS), Autism Screening Quotient (AQ-10), Autism Spectrum Screening Questionnaire-Revised Extended Version (ASSQ-REV), Developmental Behavior Checklist-Autism Screening Algorithm (DBC-ASA), Diagnostic Behavioral Assessment for Autism Spectrum Disorder-Revised (DiBAS-R), A DSM-5 teacher screening questionnaire for autism & social communication disorders (EDUTEA), the Autism Diagnostic Inventory-Telephone Screening in Spanish (ADI-TSS), Adapted Autism Behaviour Checklist (AABC) and the Mobile Autism Risk Assessment (MARA).

After evaluating the evidence for the twelve tools, two of them (SCQ and AQ-10) were selected for further review by a consensus group of Nigerian experts because the SCQ had evidence of cross-cultural use and the AQ-10 was the adolescent version. The group of experts examined the content and face validity of both tools. After the consensus group consultations in Nigeria, as

discussed in Chapter 5, (Nwokolo et al., in press), the Social Communication Questionnaire (SCQ) was selected, and adjusted slightly to contain more culturally relevant examples. The SCQ is a brief 40-item parent or caregiver screening measure used widely in research (Berument et al., 1999). Administration of the SCQ is time-efficient, requiring no costly or special training. The group agreed that the SCQ was more robust and comprehensive than the AQ-10, with questions that examined the relevant autism spectrum domains.

To establish the usefulness of the SCQ in Nigeria, this study aimed to a) validate the structure of the SCQ in the Nigerian population using confirmatory factor analysis (CFA), b) examine the internal consistency, discriminant, and convergent validity of the SCQ, c) derive an appropriate cut-off score based upon sensitivity and specificity and d) derive the positive and negative predictive values.

The recruitment of participants from the relevant centres was purposive to allow the inclusion of some persons suspected to have autism spectrum disorder and some thought not to have the disorder, given the study's aims.

7.1.1 Skill and Experience of the Research Assistants

One research assistant was involved in this study. The research assistant was a Qualified Autism Service Practitioner- Supervisor (QASP-S). The QASP-S is a certificate obtained from the Qualified Applied Behavior Analysis Credentialing Board (QABA); a US-based organisation recognised in Nigeria. The QASP-S administered the autism and intellectual disability screening tools. Study details were shared with the research assistants. The lead researcher provided one full day of training to the QASP-S on administering the consent forms, information sheets and screening tools, followed by three days of demonstration. Following this, the QASP-S was observed while implementing the procedure. No procedural errors were observed. The main researcher administered the ADOS-2 and VABS-3.

7.2 Methods

7.2.1 Design

A between-groups design was then used with two groups of participants: adolescents and young people with and without suspected autism spectrum disorder.

7.2.2 Participants

An adolescent or young person was eligible to take part in this study in the 'suspected ASD' group, if they were a) between 11- and 26 years old, b) identified by a doctor as having a clinical diagnosis of ASD, and/or c) enrolled in a special education school or a special centre and d) had a parent, guardian, or caregiver with adequate lifetime information regarding the adolescent. The 'non-autistic' participants (i.e., those not suspected of having autism) were recruited through the main researcher's contacts and social circle. The study occurred within three of Nigeria's geopolitical zones: Abuja, Enugu, and Lagos. Participants were recruited from day centres, special schools, child and adolescent mental health care services, local community organisations, religious organisations, and public advertisements.

Initially, the first ten participants recruited from the Child & Adolescent Mental Health hospital had been clinically diagnosed with autism; however, upon administering the ADOS-2 and the SCQ to the carers, differences in categorisation were observed. Following this, the main researcher informed the referring clinicians to simply send willing candidates to the team without any note as to diagnosis, so as to eliminate possible bias, ensuring the research team would be blind to the clinican's view of diagnosis.

As a result, two hundred and ten adolescents and young adults, 124 (59%) male and 86 (41%) female (M_{age} = 15.88 years; Mdn_{age} = 15.29 years; SD = 3.69; range = 10.90 – 26.96 years) took part in this study. The age distribution was grouped as follows: 11 – 13-year-olds (n = 76; 36.2%), 14 – 15-year-olds (n = 42; 20%), 16 – 17-year-olds (n = 51; 24.3%) and 18 years and above (n = 41; 19.5%). Initially, 245 adolescents and young adults were invited to participate in this study, but 35 declined to participate or did not respond to further attempts to contact them, so that finally 210 took part. Out of the 210, a further 5 participants did not complete the SCQ and were excluded from the analysis leaving 205 as the total number of participants.

7.2.3 Procedure

7.2.3.1 Ethics

A positive ethical opinion was obtained from the University of Kent, Tizard Centre Ethics Committee, the National Health Research Ethics Committee (NHREC; NHREC/01/01/2007-16/09/2019) and the Federal Neuro-Psychiatric Hospital, Yaba, Lagos, Nigeria

(FNPH/HREC/20/09). Ethics approvals are in the Appendices (1 to 4). All participants were provided with written information about participating in this study, assent, and consent forms, including easier-to-read versions (Appendix 11, 12, 13, 14, 15, 16, 17 & 18). Parental informed consent was sought for those aged under 18 years of age (Appendix 10). All participants were given the feedback form which contained information of whom to contact should they have any concerns about the research such as withdrawing (Appendix 19). Participants were encouraged to take breaks as needed during testing.

7.2.3.2 Study Design and Procedure

Each participant was seen either in their school, place of worship, centre, clinic or home and was invited to complete the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2) while the SCQ was completed by their parent, caregiver, or guardian. Participants were included in the autism group if they met the autism or autism spectrum classification on the ADOS-2. The ADOS-2 was administered to the adolescent by the first author, who has been trained to meet the ADOS-2 to clinical reliability. Video recordings of approximately 5% of the ADOS-2 administrations were independently reviewed and scored by a second rater (GM). ADOS-2 total scores between the raters correlated at 0.98 (p<.001) with a 95% confidence interval from 0.94 to 1. Following review of the ADOS-2 categorisations, into ASD/no ASD, the inter-rater agreement was k = 1. 7.2.4 Measures

7.2.4.1 Social Communication Questionnaire (SCQ)

The SCQ is a brief 40-item parent or caregiver-report screening measure modelled after the Autism Diagnostic Interview-Revised (ADI-R) and has been used widely in research (Berument et al., 1999; Rutter et al., 2003). It is a screening tool and cannot be used for the diagnosis of autism. The SCQ is designed for anyone 4 years old and above, and it takes about 10 – 15 minutes to complete and about 5 minutes to score. The measure has two versions: the lifetime and the current versions. Both focus on symptoms of autism most likely to be observed by the individual's principal caregiver, who must be familiar with the individual's developmental history and current behaviour. The lifetime version was used in this study, given the age range of the participants (11 – 26 years). In addition, Wei et al. (2015) reported that the lifetime version has better psychometric properties.

7.2.4.2 Autism Diagnostic Observation Schedule, Second Edition (ADOS-2)

The Autism Diagnostic Observation Schedule, 2nd Edition, is a very widely used tool for assessing the presence of autism (Chojnicka & Pisula, 2017; Lebersfeld, Swanson, Clesi & O'Kelley, 2021). It is a semi-structured standardised clinician tool which uses a hierarchy of presses across a range of play-based activities to observe behaviour, communication, social interaction and imaginative use of materials. An overall score is obtained with cut-offs for ASD. The ADOS-2 has five modules – toddler and modules 1 to 4. Modules 1, 2 and 3 were used in this study. To determine the applicable module for each participant, the suggested guidelines in the ADOS-2 manual was followed. The guideline includes evaluation of the individual's expressive language and determining the chronological age. The ADOS-2 takes between 60 – 90 minutes to administer and score.

According to NICE guidelines in the UK (NICE, 2017), the ADOS-2 should be used alongside the ADI-R or DISCO interview to make a certain diagnosis of autism. However, resources did not allow for the ADI-R or DISCO, so the ADOS-2 was used alone as the gold standard.

7.2.5 Data Analysis

7.2.5.1 Overview

The Statistical Package for the Social Sciences – IBM SPSS version 26 and Jeffreys's Amazing Statistics Program (JASP) version 0.16.3, an open-source statistical package, were used for analyses. Except for the confirmatory factor analysis (CFA) done with JASP, all other analyses were done with SPSS. CFA was performed to confirm the applicability and validity of the original SCQ constructs to the Nigerian adolescent population. The SCQ's performance as a screening tool was compared to the ADOS-2 classification, while correlations between the 40 SCQ items and ADOS-2 were calculated using Pearson's *r*. The discriminant and convergent validities of the SCQ were examined. Internal consistency was calculated using Cronbach's alpha. Receiver Operating Characteristic (ROC) analyses were used to calculate the Area under the Curve (AUC) to examine how well the SCQ identified participants with and without an autism spectrum disorder, with reference to sensitivity and specificity to identify optimal cut-offs. The positive (PPV) and negative predictive (NPV) values were calculated from the results.

7.2.5.2 Missing Data

5 participants did not complete the SCQ, and 6 did not complete the ADOS-2. Therefore 5 participants were excluded from the SCQ analysis and 6 from the ADOS-2 analysis.

7.2.5.3 Confirmatory Factor Analysis (CFA)

The SCQ is a standardised measure used widely in research and different languages, for example, South Africa (Bozalek, 2013), Uganda (Awadu, 2021) and China (Liu et al., 2022). Some studies have examined the structure (Uljarević et al., 2021), psychometric properties (Wei et al., 2015) and its utility as a screening tool (Chestnut, Wei, Barnard-Brak & Richman, 2017). Whereas some Nigerian professionals are conversant with the SCQ, the applicability of the SCQ has not been examined. Thus, a confirmatory factor analysis (CFA) was performed to evaluate the validity of the original SCQ constructs in the Nigerian adolescent population.

Initial model fit using the original four factors (social interaction, communication, abnormal language, and stereotypic behaviour, of Berument et al., 1999) was examined using the JASP Version 0.16.3. This was followed by a bootstrapping with replacement (based on a sample size of 5000) and the diagonal weighted least squares (DWLS) estimator, which are appropriate for small sample sizes (Mîndrilă, 2010; DiStefano & Morgan, 2014; Koğar & Koğar, 2015). The model fit was first evaluated based on the chi-square (χ^2) goodness-of-fit statistics. Due to the sample size, literature and various opinions and criticisms about the chi-square, other indices were also examined: the chi-square/*df* ratio <3, the comparative fit index (CFI) ≥ .90, the Tucker-Lewis index (TLI) ≥ .90, the root mean square error of approximation (RMSEA) ≤ .08, the goodness of fit index (GFI) ≥ .90 and the standardised root mean square residual (SRMR) between .05 and .08 (Hu & Bentler, 1999; Prudon, 2014; Newsom, 2018a, 2018b; Mîndrilă, 2010).

7.2.5.4 Internal Consistency

Cronbach's alpha, considered an adequate measure of internal consistency (Mokkink et al., 2018a; Terwee et al., 2007), was used to assess the internal consistency of the SCQ. A Cronbach's alpha greater than or equal to 0.70 is acceptable (Tavakol & Dennick, 2011).

7.2.5.5 Criterion Validity

The criterion validity of the SCQ was determined by assessing both the discriminative and convergent validities. Terwee et al. (2007) and Mokkink et al. (2018a) suggest a good correlation with the 'gold standard' tool if the correlation is ≥ 0.70 or AUC ≥ 0.70.

- Discriminative Validity. The discriminative ability of the SCQ was determined by examining the AUC.
- Convergent Validity**.** Convergent validity was assessed by examining the extent of correlation between the SCQ scores and the ADOS-2 classifications. The correlation was determined by using Pearson's correlation coefficient ***r***.

7.2.5.6 Sensitivity and Specificity

The sensitivity of the SCQ refers to the probability of it correctly identifying individuals with ASD, while the specificity refers to its probability of correctly identifying those who do not have ASD (Trevethan, 2017). The optimal cut-off score for the SCQ was based on the ROC analysis, while specificity and sensitivity were determined from the AUC (Streiner & Cairney, 2007; Lasko, Bhagwat, Zou & Ohno-Machado, 2005). Sensitivity and specificity cut-off values were guided by Glascoe (2005).

7.2.5.7 Positive Predictive Value (PPV) and Negative Predictive Value (NPV)

PPV and NPV determine those who genuinely have or do not have an ASD. The SCQ's PPV was determined using the formula [(true positives/(true positives + false positives)]*100 and NPV as [(true negatives/(true negatives + false negatives)]*100. Because both values relate to prevalence, there are no agreed cut-off values for PPV & NPV for a screening tool (Glascoe, 2005; Wong & Lim, 2011).

7.3 Results

7.3.1 Descriptive Statistics

SCQ scores for all 205 participants ranged from 0 to 30 points (M = 8.42; Mdn_{score} = 6.00; SD = 6.89). Total SCQ scores did not differentiate by sex (t_{205} = .34, p = .74). Descriptive statistics for the SCQ scores per age group are in Table 23 and the descriptive statistics for the ADOS-2 scores per age group are in Table 24. The distribution of the 204 participants with and without

ASD (as defined by the ADOS-2 autism spectrum cut off score for each module) is in Table 25. The participants distribution by age group according to the modules used in the ADOS-2 administration are in Table 26. Sixteen percent of the participants received module 1, 20% for module 2 while 64% received module 3. Out of the 64% (n=130) that the module 3 was administered to, approximately 28% (n=19) were classified as having ASD (Table 25).

Table 23 - Descriptive statistics of SCQ scores by age groups (total n=205)

	SCQ Total Score			
	11 - 13 years	14 - 15 years	16 - 17 years	18 years & above
Number	74	41	50	40
Mean	8.81	8.49	8.02	8.13
Std. Deviation	7.34	7.75	6.01	6.34
Minimum	0.00	0.00	0.00	0.00
Maximum	30.00	27.00	24.00	30.00

Table 24 - Descriptive statistics of ADOS-2 scores by age groups (total n=204)

	ADOS-2 Total Score			
	11 - 13 years	14 - 15 years	16 - 17 years	18 years & above
Number	74	40	50	40
Mean	6.01	5.45	5.52	4.99
Std. Deviation	7.12	6.38	6.76	6.35
Minimum	0.00	0.00	0.00	0.00
Maximum	22.00	20.00	23.00	20.00

Table 25 - Descriptive statistics of the ADOS-2 classification for participants with & without autism spectrum disorder based on the autism spectrum cut-off score for each module.

		ADOS-2 Classification					
		No ASD		ASD		Total	
		N	%	N	%	N	%
Age Groups	11 to 13 years old	47	34.81	27	39.13	74	36.27
	14 to 15 years old	26	19.26	14	20.29	40	19.61
	16 to 17 years old	32	23.70	18	26.09	50	24.51

	18 years old and above	30	22.22	10	14.49	40	19.61
ADOS-2 Modules	1	5	3.70	28	40.58	33	16.18
	2	19	14.07	22	31.88	41	20.10
	3	111	82.22	19	27.54	130	63.73
Gender	Female	55	40.74	29	42.03	84	41.18
	Male	80	59.26	40	57.97	120	58.82
	Total	135	100.00	69	100.00	204	100.00

Table 26 - Frequency of participants by age group according to the modules used in the ADOS-2 administration.

		Age Groups				
		11 to 13 years old	14 to 15 years old	16 to 17 years old	18 years old and above	Total
ADOS-2 Modules	1	12	6	9	6	33
	2	19	8	4	10	41
	3	43	26	37	24	130
	Total	74	40	50	40	204

7.3.2 Confirmatory Factor Analysis (CFA)

The CFA had an acceptable model fit, CFI = 1.00, TLI = 1.03, SRMR = 0.08, RMSEA = .00 and GFI = .96. Figure 2 shows the model plot and factor loadings in Table 27. The initial hypothesis was that the original four factors (social communication, social interaction, abnormal language, and stereotypic behaviour) of the SCQ may not be applicable in the Nigerian context, but it transpired that the original four factor structure could be maintained. Items 5, 9 and 13 had factor loadings that were slightly below 0.3 (0.26, 0.20 & 0.20, respectively) but were retained in the SCQ as removing was not deemed impactful to the structure of the SCQ. Overall, the factor loadings indicated that all the items could be retained in the SCQ.

7.3.3 Internal Consistency

The Cronbach's α = .88 for the total sample and α = .86 for the ASD group on all original four domains of the SCQ indicated adequate internal consistency of the SCQ items, while for the non-ASD group α = .59. In the entire sample, Cronbach's α = .85 for the social communication and interaction domain and .66 for the restricted, repetitive and stereotypic pattern of behaviours domains (RRSB) – items 7, 8, 11, 12, 13, 14, 15 & 16. With the addition of the self-injurious items (17, 18 & 38) to the RRSB, Cronbach's α = .71.

7.3.4 Criterion Validity

7.3.4.1 Discriminative Validity

The cut-off score of 10 on the SCQ showed that the SCQ could differentiate between those with and without ASD, (using the cut-off score for autism spectrum applicable for each ADOS-2 module). An AUC of 1 would indicate a perfect screening tool. At the cut-off score of 10, the AUC was .83.

7.3.4.2 Convergent Validity

Overall convergent validity was indicated by a significant Pearson's correlation between the total SCQ scores and ADOS-2 classifications for the 204 participants, $r = .70$, $p < 0.001$, showing a strong correlation and effect size (Cohen, 1992).

Table 27 - Factor Loadings

				95% Confidence Interval				
Factor	**Indicator**	**Estimate**	**SE**	**Lower**	**Upper**	**Z**	**p**	**Stand. Estimate**
Social Interaction	SCQ_Item17	0.1551	0.0254	0.1054	0.2049	6.11	< .001	0.427
	SCQ_Item19	0.1775	0.0330	0.1129	0.2422	5.39	< .001	0.380
	SCQ_Item21	0.2476	0.0317	0.1855	0.3098	7.81	< .001	0.529
	SCQ_Item22	0.2098	0.0622	0.0879	0.3317	3.37	< .001	0.244
	SCQ_Item26	0.1936	0.0252	0.1443	0.2429	7.69	< .001	0.523
	SCQ_Item27	0.0993	0.0158	0.0684	0.1302	6.30	< .001	0.439

Table 27 - Factor Loadings

Factor	Indicator	Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
	SCQ_Item28	0.2460	0.0248	0.1975	0.2946	9.93	< .001	0.648
	SCQ_Item29	0.2195	0.0293	0.1622	0.2768	7.50	< .001	0.512
	SCQ_Item30	0.2088	0.0243	0.1612	0.2564	8.59	< .001	0.577
	SCQ_Item31	0.2220	0.0289	0.1653	0.2787	7.68	< .001	0.522
	SCQ_Item32	0.1520	0.0236	0.1057	0.1983	6.43	< .001	0.447
	SCQ_Item33	0.2132	0.0203	0.1733	0.2530	10.49	< .001	0.673
	SCQ_Item34	0.3217	0.0250	0.2727	0.3708	12.86	< .001	0.782
	SCQ_Item35	0.3335	0.0295	0.2757	0.3914	11.30	< .001	0.714
	SCQ_Item36	0.2423	0.0234	0.1964	0.2882	10.35	< .001	0.665
	SCQ_Item37	0.1669	0.0195	0.1288	0.2050	8.58	< .001	0.573
	SCQ_Item38	0.1571	0.0240	0.1101	0.2040	6.56	< .001	0.455
	SCQ_Item39	0.3125	0.0294	0.2549	0.3702	10.63	< .001	0.682
	SCQ_Item40	0.2389	0.0240	0.1919	0.2858	9.97	< .001	0.646
Communication	SCQ_Item2	0.0574	0.0152	0.0276	0.0871	3.77	< .001	0.304
	SCQ_Item9	0.0688	0.0274	0.0151	0.1226	2.51	0.012	0.187
	SCQ_Item15	0.1451	0.0245	0.0971	0.1931	5.93	< .001	0.420
	SCQ_Item20	0.1711	0.0292	0.1139	0.2284	5.86	< .001	0.416
	SCQ_Item24	0.3350	0.0220	0.2920	0.3780	15.26	< .001	0.885
	SCQ_Item25	0.3412	0.0213	0.2994	0.3830	16.00	< .001	0.912
Abnormal Language	SCQ_Item3	0.3014	0.0396	0.2238	0.3790	7.61	< .001	0.619
	SCQ_Item4	0.2010	0.0375	0.1276	0.2745	5.37	< .001	0.459
	SCQ_Item5	0.1284	0.0452	0.0398	0.2169	2.84	0.004	0.257
	SCQ_Item6	0.1757	0.0375	0.1022	0.2491	4.69	< .001	0.400
	SCQ_Item7	0.3082	0.0375	0.2347	0.3816	8.22	< .001	0.668
Stereotypic Behaviour	SCQ_Item8	0.1742	0.0397	0.0964	0.2521	4.39	< .001	0.362
	SCQ_Item10	0.2307	0.0332	0.1657	0.2958	6.95	< .001	0.534
	SCQ_Item11	0.2676	0.0317	0.2054	0.3297	8.43	< .001	0.640
	SCQ_Item12	0.2048	0.0329	0.1403	0.2694	6.22	< .001	0.485
	SCQ_Item13	0.1048	0.0408	0.0247	0.1848	2.56	0.010	0.212

Table 27 - Factor Loadings

Factor	Indicator	Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
	SCQ_Item14	0.1500	0.0338	0.0839	0.2162	4.44	< .001	0.353
	SCQ_Item16	0.1426	0.0272	0.0892	0.1960	5.24	< .001	0.421
	SCQ_Item18	0.1816	0.0324	0.1181	0.2452	5.60	< .001	0.449

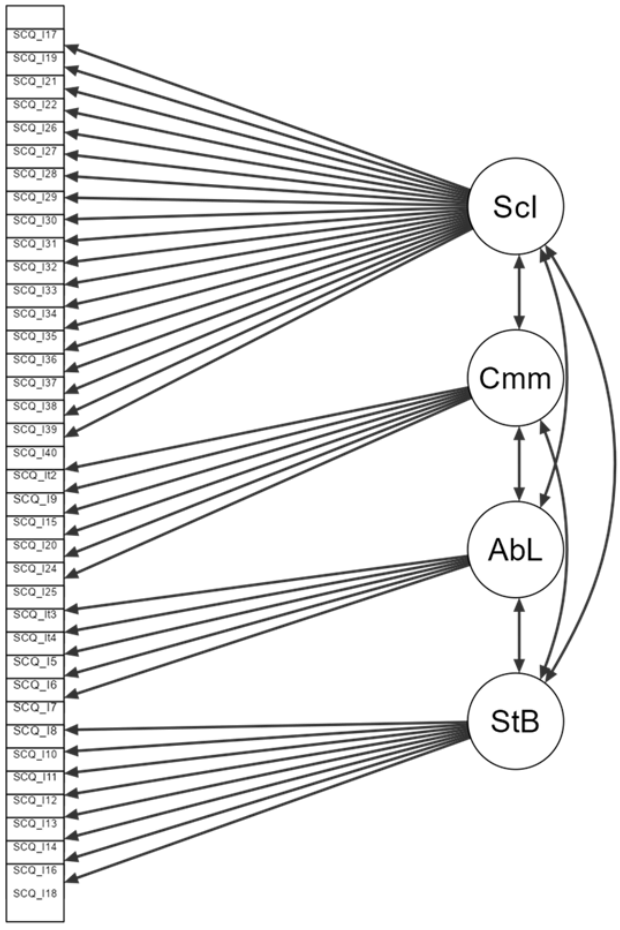


Figure 2: Model plot

7.3.4 Sensitivity and Specificity

At the recommended cut-off score of 15, ROC analysis revealed an overall AUC for the 204 participants as .76, $p < 0.001$, 95% CI [.68, .84] with PPV = .54, NPV = .99, sensitivity = .95 and specificity = .81. While the specificity and sensitivity met the minimum requirements, the PPV did not meet the minimum standard (Glascoe, 2005). Lower cut-offs were explored by stepwise reduction to determine the best cut-off score. Lowering the cut-off score to 10, 11 and 12 improved the values; however, a cut-off score of 10 gave the best result overall, as shown in Table 28. With a cut-off score of 10, sensitivity = .81 and specificity = .88, applicable to the entire population.

For the specific age groups, using the cut-off score of 10, the following results were obtained: 11 – 13-year-olds, AUC = .83, $p < 0.001$, 95% CI [.77, .90], $N = 74$, 14 – 15-year-olds, AUC = .84, $p < 0.001$, 95% CI [.94, .99], $N = 40$, 16 – 17-year-olds, AUC = .84, $p < 0.001$, 95% CI [.71, .97], $N = 50$, 18 years and above, AUC = .83, $p < 0.001$, 95% CI [.67, 1.00], $N = 40$. Sensitivity and specificity with optimal cut-off scores for each group are shown in Table 29. Additionally, sensitivities and specificities for each SCQ cut-off score (10, 11, 12 and 15) were explored for the different ADOS-2 modules. At the recommended cut-off score of 15, Module 1 had a sensitivity of 1 and specificity of .33, Module 2 a sensitivity of 1 and specificity of .66 while Module 3 had a sensitivity of 0 and specificity of 1. At the cut-off score of 10, Module 1 had a sensitivity of 1 and specificity of .56, Module 2 a sensitivity of .71 and specificity of .65 while Module 3 had a sensitivity of .68 and specificity of .87 and these figures are shown in Table 30. Some of these values which are low, are due to the low number of participants for each module relative to the total population values based on 204 participants. Based on available data, this study is the first to explore the usefulness of differentiated cut-off scores per age group for the SCQ. From this study, a cut-off score of 10 is preferable for all those under 18yrs, and a cut-off score of 12 is best suited for participants aged 18 and above, as all the psychometric properties met the minimum standard (Table 29).

7.3.5 Positive Predictive Value (PPV) and Negative Predictive Value (NPV)

Using the SCQ cut-off score of 10, PPV = .75 and NPV = .91, showing that the SCQ can correctly identify those with ASD and those without ASD. The PPV and NPV for the different

ADOS-2 modules were also explored with PPV ranging between .53 and 1 while the NPV ranged between .68 and 1. The results are in Table 30.

7.4 Discussion and Conclusion

With the increased awareness of autism spectrum disorder in Nigeria, parents of younger children now seek screening and early intervention. However, the older children and adolescents who missed early screening and diagnosis need to be known. To detect ASD in these adolescents, a validated and easy-to-use screening tool is required. The SCQ was identified through a systematic review (Nwokolo et al., 2022) and agreed on as appropriate by a focus group (Nwokolo et al., in press). Thus, the goals of this study were to a) validate the structure of the SCQ in the Nigerian population using confirmatory factor analysis (CFA), b) examine the internal consistency, discriminative, and convergent validity of the SCQ, c) derive an appropriate cut-off score based upon sensitivity and specificity and d) derive the positive and negative predictive values.

7.4.1 Confirmatory Factor Analysis (CFA)

While the SCQ is an established and widely used measure in both research and clinical settings, the accuracy and psychometrics of the SCQ have mainly been examined in North and South America, Europe, and Australia (Chestnut et al., 2017). Studies confirming its appropriateness in the African context, especially among adolescents, were non-existent. However, studies which examined the discriminative validity in young and older children aged between 2.5 and 17 years in South Africa (Bozalek, 2013) and Uganda (Awadu, 2021) were found. Based on existing literature, the scarcity of cross-cultural research in the African context and the aims of this study, a CFA was done. The CFA results revealed that the Nigerian population could retain the original four-factor structure, bearing in mind the limitations of the sample size.

7.4.2 Internal Consistency

The SCQ's internal consistency was adequate, with a Cronbach's alpha of .88, indicating the tool's ability to capture the concept of autism spectrum disorder.

7.4.3 Criterion Validity

Although at the published cut-off score of 15, the sensitivity, specificity and NPV met recommended criteria (Glascoe, 2005), the PPV was not optimal. With a reduction in the cut-off score to 10, all the properties met the minimum criteria, with the SCQ adequately discriminating between those with ASD and those without ASD. There was a strong positive relationship between the SCQ scores and group classification (with or without ASD), which showed that as the SCQ scores increased, the more likely an individual would have an ASD. The SCQ, as a screening tool, correlated well with the ADOS-2 (r = .70), showing that the SCQ is a valid screening instrument for use with the Nigerian adolescent population.

7.4.4 Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value

Based on the results of the ROC analyses, the sensitivity, specificity, PPV and NPV met the acceptable criteria of 70% for screening tools (Glascoe, 2005). These results are based solely on the ADOS-2 ratings and classifications from this research, since participants were not formally diagnosed. Only the ADOS-2 scores were used to classify those who belonged to the ASD and no-ASD groups. Initially, at the recommended cut-off score of 15, PPV = .54, NPV = .99, sensitivity = .95 and specificity = .81. However, the discriminating ability improved by reducing the cut-off to 10, giving a sensitivity of .81 and specificity of .88. This cut-off score of 10 is like the results obtained by Bozalek (2013) in the South African sample (cut-off = 10, sensitivity = 1, specificity = .95) and Awadu (2021) in Uganda (cut-off = 10, sensitivity = 1, specificity = .93). Other studies (Kim et al., 2015; Snow & Lecavalier, 2008; Schanding, Nowell & Goin-Kochel, 2012) also recommended a reduction in the cut-off score from 15 for better outcomes. In assigning the cut-off, the distinctions between sensitivity and PPV and between specificity and NPV in a screening and clinical context were considered (Trevethan, 2017; Akobeng, 2007). Classifying participants solely on sensitivity and specificity values differs from classifying them in combination with the PPV and NPV. PPV and NPV are influenced by the prevalence and depend on the population being investigated. Participants identified by medical professionals as autistic were sampled in this study; as such, the PPV and NPV were considered in addition to sensitivity and specificity to determine the best cut-off score. Since the Lifetime version of the SCQ was used, it is possible that some of the respondents of the older participants may not have

an absolute recollection of the early years of their wards. For this reason, the scores will be affected, and a lower cut-off ensures that persons who may have autism are not missed. Should the sample age in any study be homogenous, which is highly unlikely, specific cut-offs are recommended for the different age groups, as shown in Table 29. Overall, the results showed that the SCQ correctly identified adolescents with and without ASD.

7.5 Limitations

There are some limitations to this study. First, the sample size was relatively small for CFA, although our model was associated with a good fit. Secondly, the participants were mainly from urban settings and had good literacy skills; thus, it cannot be assumed that the psychometric properties will be the same when used in rural settings, where questions may need to be read to respondents. Thirdly, while we recognise that the use of the original English version of the SCQ may be judged insufficient for a validation study in an ethnic and culturally diverse setting as Nigeria, English is the official language in Nigeria, and we ensured the examples given were culturally appropriate. English as the official language, or pidgin (a variation of English) is widely spoken by most people especially the urban dwellers. Additionally, while urban populations may be similar, given that there is insufficient evidence of formal validation of any autism screening tool for the Nigerian population, validation of English SCQ was deemed a viable start. Further studies to explore the translation and validation of the SCQ in the three major languages (Hausa, Igbo and Yoruba) is recommended. Fourthly, participants were categorised as autistic based upon the ADOS-2 only and I did not undertake an additional assessment such as the ADI-R, DISCO or generate a thorough developmental history. A similar criticism about the use of English may arise concerning the use of the ADOS-2, however, for the same reasons that the English SCQ was used, and in the absence of other available tools, the ADOS-2 was deemed appropriate for use with the Nigerian urban population. It is possible that in doing so, functional and stimulus biases may have been introduced in that the Nigerian participants may not have been offered the same opportunity to demonstrate knowledge while eliciting the intended response as participants in the original ADOS-2 study. The use of ADOS-2 materials ‘as is’ without any cultural adaptations or modifications, has been supported in a study in South Africa (Smith, Malcolm-Smith & de Vries, 2017). Smith et al. (2017), assessed the language, social activities, and materials of the translated Afrikaans version of the ADOS-2 for

participants who they referred to as ‘coloured’, from the Western Cape in South Africa. They found that the participants were sufficiently familiar with the items and materials. However, to assist those assessing the participants and to improve cultural sensitivity Smith et al. (2017) generated guidelines in Afrikaans for those who were administering the ADOS-2. When tools are translated, uniqueness in language must be considered (Smith et al., 2017; International Test Commission, 2017). However, the ADOS-2 was not translated for participants for this present study since it was used within urban settings where familiarity with the materials and language were not a problem. A study examining the validity of the ADOS-2 in the Nigerian context in more detail is recommended. Despite these limitations, the SCQ appears to be a useful screening tool for ASD in Nigerian adolescents.

7.6 Conclusion

In conclusion, the SCQ Lifetime form’s psychometric properties met acceptable screening tools standards across the entire sample and all age groups of Nigerian adolescents and young people. All items of the SCQ Lifetime version are relevant, with culturally relevant examples used as applicable. Based on available data, this study is the first to explore the usefulness of differentiated cut-off scores per age group for the SCQ. From this study, a cut-off score of 10 is recommended for all those under 18yrs, and a cut-off score of 12 for participants aged 18 and above, as all the psychometric properties met the minimum standard. Further studies exploring these cut-offs are recommended. The SCQ Lifetime form can be used as a screening tool for identifying Nigerian adolescents likely to have autism spectrum disorder and help ensure referrals for further diagnosis. Using the suggested cut-offs for specific age groups will be beneficial in clinical settings.

Table 28 - Sensitivity and specificity for the various potential cut-off scores of the SCQ

Cut-off Score	PPV	NPV	Sensitivity	Specificity	AUC	Lower	Upper
10	0.75	0.91	0.81	0.88	0.83	0.77	0.90
11	0.68	0.94	0.85	0.85	0.81	0.74	0.88
12	0.65	0.96	0.88	0.84	0.80	0.73	0.88
15	0.54	0.99	0.95	0.81	0.76	0.68	0.84

Table 29 - Sensitivity and specificity for the various potential cut-off scores of the SCQ per age group

	SCQ Cut-off Score >10					SCQ Cut-off Score >11					SCQ Cut-off Score >12				
	Total (n=204)	11 - 13 years (n = 74)	14 - 15 years (n = 40)	16 - 17 years (n = 50)	18 years & above (n = 40)	Total (n=204)	11 - 13 years (n = 74)	14 - 15 years (n = 40)	16 - 17 years (n = 50)	18 years & above (n = 40)	Total (n=204)	11 - 13 years (n = 74)	14 - 15 years (n = 40)	16 - 17 years (n = 50)	18 years & above (n = 40)
PPV	**0.75**	0.74	0.71	0.78	0.80	0.68	0.67	0.57	0.72	0.80	0.65	0.63	0.57	0.67	**0.80**
NPV	**0.91**	0.91	0.96	0.91	0.87	0.94	0.94	0.96	0.94	0.93	0.96	0.96	0.96	0.97	**0.93**
Sensitivity	**0.81**	0.83	0.91	0.82	0.67	0.85	0.86	0.89	0.87	0.80	0.88	0.89	0.89	0.92	**0.80**
Specificity	**0.88**	0.86	0.86	0.88	0.93	0.85	0.83	0.81	0.86	0.93	0.84	0.82	0.81	0.84	**0.93**
AUC ($p < 0.001$, 95% CI)	0.83	0.83	0.84	0.84	0.83	0.81	0.80	0.77	0.83	0.87	0.80	0.79	0.77	0.82	0.87
Lower	0.77	0.72	0.94	0.71	0.67	0.74	0.68	0.59	0.70	0.71	0.73	0.67	0.59	0.68	0.71
Upper	0.90	0.69	0.99	0.97	1.00	0.88	0.92	0.94	0.97	1.02	0.88	0.91	0.94	0.96	1.02

Table 30 – Sensitivity, specificity, PPV and NPV for the potential cut-off scores of the SCQ by ADOS-2 modules

	Module 1				Module 2				Module 3			
	SCQ >10	SCQ >11	SCQ >12	SCQ >15	SCQ >10	SCQ >11	SCQ >12	SCQ >15	SCQ >10	SCQ >11	SCQ >12	SCQ >15
PPV	**0.86**	0.79	0.71	0.64	0.68	0.68	**0.68**	0.55	0.68	0.53	**0.53**	1.00
NPV	**1.00**	1.00	1.00	1.00	0.68	0.79	**0.84**	1.00	0.87	0.96	**0.97**	0.93
Sensitivity	**1.00**	1.00	1.00	1.00	0.71	0.79	**0.83**	1.00	0.68	0.71	**0.77**	0.00
Specificity	**0.56**	0.45	0.38	0.33	0.65	0.68	**0.70**	0.66	0.87	0.93	**0.92**	1.00
AUC ($p < 0.001$, 95% CI)	**0.93**	0.89	0.86	0.82	0.68	0.74	**0.76**	0.77	0.82	0.75	**0.75**	0.68
Lower	**0.84**	0.78	0.73	0.68	0.52	0.58	**0.61**	0.63	0.69	0.60	**0.60**	0.52
Upper	**1.02**	1.00	0.99	0.97	0.85	0.89	**0.91**	0.92	0.94	0.89	**0.90**	0.83

Chapter 8. Discussion

8.1 The Research Overview

Screening for and diagnosing autism and intellectual disability remains a challenge in Africa and Nigeria specifically. Attempts have been made to ascertain the burden of these disabilities in Nigeria (Oshodi et al., 2016) however, focus has been on mainly early years. Individuals who may be autistic or have intellectual disability are often noticed much later around the age of adolescence when they transition out of the close family unit. There are no tools to screen these individuals due to the lack of validated screening tools. Several studies have investigated the availability of screening tools for use in low to-middle income economies but focused on early years and young children (de Vries, 2016; Ruparelia et al., 2016; Stephens, 2012; Marlow et al., 2019). There remains a gap in research for adolescents, especially in Nigeria. Thus, the focus of this thesis was on adolescents in Nigeria. The main research questions were:

3. Is there a screening tool for intellectual disability which requires little or no training that can be used amongst Nigerian adolescents?
4. Is there a screening tool for autism which requires little or no training that can be used amongst Nigerian adolescents?

Based on these questions, this thesis aimed to identify and evaluate screening tools for autism and intellectual disability which require little or no training that can be used amongst Nigerian adolescents.

To identify the most appropriate tools, this research was broken down into several more specific studies. The findings from studies 1 and 2 culminated in validating the SCQ and SCIL amongst Nigerian adolescents (Study 3a and 3b). This chapter will discuss the results and implications for practice and research, as well as the limitations of the research and recommendations.

8.2 Overview of Studies

The systematic review aimed to a) describe and critically appraise short screening tools for the detection of intellectual disabilities and autism in children and young people aged 11 to 26 years, b) consider the psychometric properties of these tools, and c) consider the appropriateness of

using these tools across a range of cultures. Therefore, the review was conducted following the PRISMA guidelines (Page et al., 2021) and quality appraisal of identified studies was based on the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) risk of bias checklist (Mokkink et al., 2018b; Prinsen et al., 2018; Terwee et al., 2018b) and the manual as guides. As a result, 41 studies were included in the review.

Out of the 15,369 participants involved in the autism and intellectual disability studies, about 40% were between 12 and 75 years old, the remaining participants being under 12years of age. Six screening tools for intellectual disability and twelve for autism were identified, most for either young children or adults. However, adolescent-specific studies remained rare as only the studies with the AQ-10 adolescent version and the adolescent SCIL were identified. This further supported the gap identified in the literature regarding adolescent studies (Marlow et al., 2019). It is important to note that some studies that included participants aged 16 years and above treated them as adults. This may be primarily due to the screening and comparison tools, such as the SCQ, PDD-MRS and DiBAS-R, which are broad-age tools (and the WAIS is used for persons from 17 years and above). Regarding the cross-cultural adaptation and validity, most of the identified studies were conducted in the UK (34%) and the US (22%), with the remaining 44% spread across other European countries. Of the 41 studies identified, only 4 studies were conducted outside of the West (Argentina, Singapore, Turkey & Qatar), but predominantly with younger children. Similar findings were reported by Soto et al. (2015), i.e., that most available tools have been validated for use in the West and not in Africa. Most of the studies reviewed reported the sensitivity and specificity of the tools, with a few including the PPV and NPV, though none of these for use in Africa.

Selecting suitable and culturally sensitive measures was crucial to adapting any screening tools for use in Nigeria. Following the identification of possible screening tools for use with the Nigerian adolescent population, a focus group meeting was conducted aimed at considering the face, content, and cultural validities of the chosen screening tools and making adaptation recommendations for their use with Nigerian adolescents. Three themes were identified following data analysis of the focus group meeting: language, cultural relevance, and face validity. The group discussion highlighted areas around language, especially the Nigerian parlance, use of and meaning of words. For instance, one of the members commented that there

was a ‘Nigerian English’, implying that the use of the language and the expected response to the spoken language would be different from the expectation of a typical Westerner. An example was the word ‘ritual’ in the SCQ. Given the religious and social models of disability of Nigerians and the cultural context, the group iterated the potential misconception of the word. The group suggested using an alternative word without changing the question’s intent. Additionally, more contextualised examples were included alongside those in the questions for the SCQ. Regarding the SCIL, simple modifications such as changing ‘GP’ to ‘doctor’, removing the Euro symbol to make the numbers more generic, and including other levels of education in question 2, were discussed, both with the Dutch originators of the SCIL (as there was no existing English version) as well as the Nigerian focus group. Since the modifications proposed for both the SCIL and SCQ were minimal and did not remove from the intended outcome, the focus group’s suggestions were accepted.

Following the recommendation of the focus group, the SCIL was validated with the Nigerian adolescents. The psychometric properties of the screening tool were assessed. The aim was to examine a) the component structure of the SCIL, b) the internal consistency, discriminant, and convergent validity of the SCIL, c) the likely appropriate cut-off score based upon sensitivity and specificity, and d) the positive and negative predictive values. A total of 209 participants were assessed with the SCIL. Dimension reduction was considered using principal components analysis. Discriminative and convergent validity were examined, along with the sensitivity, specificity, PPV and NPV of the SCIL in identifying those with intellectual disabilities. The SCIL had good internal consistency, discriminative, and convergent validity. Dimension reduction was not necessary. A cut-off score of 10 revealed sensitivity = .74, specificity = .96, PPV = .92 and NPV = .86 for identifying those with an intellectual disability. AUC was .88. The psychometric properties were excellent, with a large effect size and met the suggested standards for screening tools (Glascoe, 2005), thereby suggesting that the SCIL can be used among Nigerian adolescents.

A validation study was also conducted with the SCQ. The parents and caregivers of two hundred and five adolescents completed the SCQ Lifetime form while the adolescents were assessed for ASD using the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2). The aims of the study were to a) validate the structure of the SCQ in the Nigerian population using

confirmatory factor analysis (CFA), b) examine the internal consistency, discriminative, and convergent validity of the SCQ, c) derive an appropriate cut-off score based upon sensitivity and specificity and d) derive the positive and negative predictive values. Convergent and discriminative validity were examined, along with the sensitivity and specificity of the SCQ in identifying participants with an autism spectrum disorder. Results showed that the SCQ had good internal consistency, discriminant, and convergent validity. A cut-off score of 10 revealed sensitivity = .81 and specificity = .88 for identifying autism spectrum disorder. AUC was .83. The results of the study provide evidence to support the retention of the original four factors of the SCQ Lifetime form.

8.3 Implications for Practice

There are several implications of the systematic review findings for practice. First is the scarce information on adolescent-specific screening tools for autism or intellectual disability. While the gains from early identification of individuals with autism or intellectual disability are important, persons who missed early screening will possibly remain unidentified. This also means that practitioners will continue to focus on the early years to the detriment of adolescents, especially in Africa and Nigeria specifically where early screening is at its infancy. Secondly, as highlighted in Chapter 1, developmental milestones and presentation of symptoms vary; therefore, individuals who present with signs of autism or intellectual disability later in life may not be adequately assessed without screening tools developed specifically for them. The review identified four broad-age screening tools for identifying autism (the AQ-10 and the SCQ) and for intellectual disability (the CAIDS-Q and the SCIL), however, relative to the screening tools for early years, practitioners may require more tools to be made available.

The focus group study showed the need to include relevant experts in the adaptation process of any tool for use within the African context. Since the English version of the SCIL was new, it was particularly important to include relevant professionals in the focus group to ensure comprehensiveness and relevance were assessed. Comprehensiveness is a key parameter included in the COSMIN Risk of Bias protocol (Prinsen et al., 2018, Terwee et al., 2018b; Mokinkk et al., 2018b). Similarly, the inclusion of a parent in the focus group ensured client's comprehensibility was assessed. Incorporating end-user voices in the research process allowed for the quick identification of problems and the finding of solutions. Although the SCQ is a

widely used tool, there is no record of its validation in Nigeria. Therefore, it was equally important to have input from the experts as some of the professionals in the focus group used the SCQ, and their voices were critical.

The SCIL's validation study provided preliminary support for the reliability and validity of the SCIL, showing sufficient results over several measurement properties. In a practical sense, the results of the initial validation suggest that the SCIL can be used by professionals, persons in the public health sector and primary caregivers to screen for intellectual disability reliably and accurately.

The SCQ cut-off score of 15, as recommended in the SCQ manual, yielded a sensitivity of .95 and a specificity of .81. The NPV (.99) met recommended criteria, while the PPV (.54) was below the recommended scores. Therefore, the cut-off score was adjusted to 10 in the Nigerian sample. At the adjusted cut-off score for the Nigerian sample, sensitivity was .81, and specificity was .88. In some of the other cross-cultural validation studies with a cut-off score of 15, sensitivity was between .77 and .96, while specificity was between .95 and 1. In the South African study, sensitivity = .77 and specificity = 1 (Bozalek, 2013), Uganda – sensitivity = .96 and specificity = .95 (Awadu, 2021), and Arabic – sensitivity = .80 and specificity = .97 (Aldosari et al., 2019). When cut-off scores were adjusted to between 10 and 13 in the South African, Uganda, and Arabic studies, sensitivity and specificity values improved and were between .90 and 1 for sensitivity and .80 and 1 for specificity. Thus, sensitivity and specificity values vary considerably among SCQ cross-cultural validation studies. The adjustment of cut-off scores may also reflect the subjective nature of responses on the SCQ based on the different models of disability highlighted in Chapter 1. Most of the cross-cultural validation studies cited above, except Awadu (2021), assessed the convergent validity of the SCQ against the ADOS-2. The ADOS-2 has been criticised as not being culturally sensitive (Abubakar, de Vries, & Newton, 2014; Brown-Chidsey, 2005; Deno, 2005). Abubakar et al. (2014) suggested that some of the stimuli, such as 'TV', 'spoon', and 'sink', may be unfamiliar to the participants. While this argument will hold in extremely rural areas where such items are not used, more culturally relevant stimuli can be used without changing the essence of the assessment, as has been done in other studies (Smith, 2015; Chojnicka & Pisula, 2017). They also suggested that certain behaviours like 'maintaining eye contact' were culturally inappropriate and should not be

assessed in the African context. The counter argument is that eye aversion is a learnt behaviour, as maintaining eye contact is a phylogenetic behaviour in every human and some animal organism (Farroni et al., 2002; Savalli, Resende & Gaunet, 2016). Therefore, assessing 'lack of eye contact' as part of autistic traits is relevant in the African context.

8.4 Implication for and Direction for Research

The systematic review uncovered the paucity of research on screening tools involving adolescents. While studies involving adolescents are beginning to emerge in the West, similar efforts are minimal in Africa and specifically Nigeria, the focus of this thesis. In the Nigerian context, the beliefs and orientation of most Nigerians have not allowed individuals who may require help to access professionals until they are much older. For instance, it is not uncommon to see families who bring their children for assessments or diagnosis when they are entering or have entered secondary schools. At this point, they are usually around 11 years old. Screening tools identified through this systematic review, can be used by professionals to assess these older children and adolescents. By conducting research which includes the adolescent population, an opportunity would arise for prevalence and incidence data to be gathered, thus reducing the uncertainty of estimated figures shared on the global burden of autism and intellectual disability about Nigeria. Based on the scant information on adolescent studies, there is a global opportunity for more studies. Additionally, the limited availability of adolescent-specific screening tools presents researchers with the chance to develop more autism and intellectual disability screening tools. In the African context, with limited tools, existing tools can be adapted for use with this specific population. In this case, more adaptation and cross-cultural validation studies are required. Also, these screening tools for autism or intellectual disability identified through this systematic review can be adapted for use within other African contexts.

The results of the focus group buttressed the need to include experts in the adaptation of tools outside of the original environment. Focus groups are vital to the validation process and should be used by other persons who intend to validate existing measures for use within the African context. Future use of focus groups should pay attention to contextualised relevance.

The SCIL items reflected the current definition and diagnostic criteria for intellectual disability (WHO, 2020), as mentioned in Chapter 1. This study added to the literature on assessments for

intellectual disability in Africa. It also contributed to closing the gaps around the lack of adolescent-specific research and unavailable screening tools for use in Africa identified through the systematic review. At the time of this study, there had been no formal attempt to validate existing or established intellectual disability screening tools in Nigeria. Given that presently, there is no available measure for screening intellectual disability in Nigeria, validating the SCIL for use in the Nigerian environment was an attempt to make one available. Additionally, more studies utilising the SCIL in different parts of Nigeria and its use in other English-speaking environments should be investigated to compare results.

Few autism spectrum disorder screening tools have been validated in Africa. Where validation attempts have been made, the focus was on younger children (Bozalek, 2013; Awadu, 2021; Sangare et al., 2019; Mohamed et al., 2016; Oshodi et al., 2016). Recently, a Nigerian Autism Screening Questionnaire (NASQ) was developed, but assessing convergent validity remained a challenge (Bakare et al., 2022) due to the lack of resources, as described in Chapter 2. This study, however, is the first to include a more robust psychometric evaluation. The similarity in cut-off scores (10) from this validation study and that of Bozalek (2013) conducted in South Africa may be indicative of a similar methodology in the determination of the convergent validity, unlike Awadu (2021) that retained the cut-off of 15 by comparing the SCQ to the Ten Questions which is not a diagnostic tool, in the absence of a referenced gold standard tool such as the ADOS-2. The discontinuation of the Ten Questions as a screening tool has been suggested due to its inappropriateness for detecting more complex and hidden disabilities such as mild intellectual disabilities or autism (Durkin, 2001; Olusanya & Okolo, 2006). The findings from this study provide a foundation to explore the viability of utilising the SCQ, a validated and widely used tool, in the Nigerian context with more relevant cultural modifications.

8.5 Strengths and Weaknesses

There are a number of strengths to this thesis. Firstly, there have been systematic reviews of screening tools for use in LMICs, but most have focused on very young children. The systematic review in this thesis is one of the few to focus on adolescents and Africa specifically. The systematic review identified and examined existing screening tools for autism and intellectual disability applicable to adolescents; by employing a rigorous risk of bias tool – COSMIN. The review also provided information on the psychometric properties of autism and intellectual

screening tools applicable to adolescents. Previous findings (Hirota et al., 2018; Ruparelia et al., 2016) on the scarce research with adolescent samples and insufficient studies out of Africa were corroborated through this review. Based on available data, this review, is the first to comprehensively review screening tools for autism and intellectual disability in adolescents for use in Africa. A possible weakness of the study was limiting the focus to adolescents, especially as research on autism and intellectual disabilities in Africa is still in its infancy. Widening the scope may have revealed more tools that could be adapted and validated for use in Africa in younger children.

Secondly, this is the first study to utilise the focus group technique to adapt standardised screening tools for autism and intellectual disability in Nigeria. Using the group ensured that cultural relevance was adequately explored, and all other relevant elements of language and content captured (International Test Commission, 2017). The qualitative input from the group was useful in ensuring face and content validity. Because the main researcher did not develop the screening tools (i.e., they had already been developed), the time spent at the meeting was fairly short. However, where the tool is developed from inception, using a focus group would usually not be quick or cheap.

Thirdly, the SCIL, which is a relatively new tool, was developed in the Netherlands. This study is the first to validate the translated English version in Africa, and Nigeria specifically. Fourthly, this is also the first study to examine and validate any standardised screening tool for intellectual disability in Nigeria. There is a high need for a reliable screening tool for intellectual disability in Africa and Nigeria specifically, which this study addresses. Due to the nature of the study – pioneering and validating, the opportunity to include other items which may be relevant in the Nigerian context was not examined. Additionally, a secondary strength was utilising the SCIL in a sample constituted of persons with autism and comorbid intellectual disability. The psychometric results showed that the SCIL can detect intellectual disability in adolescents with autism as well as without autism.

Fifthly, this was the first study to formally validate a standardised autism screening tool developed in the West for use in Nigeria. Sixth, it is also the first adolescent-specific screening and validation study of autism and intellectual disability in Nigeria. Seventh, this study also

added to the growing body of research in Africa regarding autism assessments and intervention. Having a valid and reliable tool for detecting autism in Nigerian adolescents could increase professionalism around screening and enable accurate data gathering. During this research, a Nigerian Autism Screening Questionnaire (NASQ) was developed and published (Bakare et al., 2022), which may seem to diminish the impact of this validation study. However, NASQ is still in its infancy and, unlike this study, does not have convergent validity. By validating the SCQ in Nigeria, the tool can be formally used by professionals with a referenced cut-off score. There is an urgent need for validated autism and intellectual disability screening tools in Nigeria, and this study has provided both. Eighth, by simultaneously validating both tools, cost and time were effectively utilised. Additionally, the simultaneous psychometric evaluation of the two different screening tools with the same participants allowed the effective use of time and the ability to examine the performance of both tools in persons with comorbid autism and intellectual disability. Ninth, the age range of participants included in the study was wide and accommodated by the broad age range (11 to 26 years) for the SCQ and SCIL. Tenth, involving experienced psychologists and research assistants knowledgeable about autism and intellectual disability who worked in the field, helped provide rapport with the parents and went some way towards neutralising any power dynamics. The eleventh is the sample size of 209 participants for the SCIL study and 205 for the SCQ, which is considered large, given the interruption caused by the COVID-19 pandemic. Twelfth, based on available data, this study is the first to explore the usefulness of differentiated cut-off scores per age group for the SCQ. From this study, a cut-off score of 10 is recommended for all those under 18yrs, and a cut-off score of 12 for participants aged 18 and above, as all the psychometric properties met the minimum standard. Finally, utilising the ADOS-2 afforded the opportunity to examine the concerns raised about the cultural relevance of the test items. Some of the issues raised about the ADOS-2 are the item, instrument, and administrative bias (Abubakar et al., 2014). Conducting the study in the urban areas eliminated these issues. However, where the tool is used in more rural areas, the issues raised may be valid. These will be addressed under the limitations of the study.

8.6 Limitations

Despite the outlined studies' strengths, there are limitations to the studies. Limitations of the individual studies have been described in each of the study chapters (4, 5, 6 & 7). The studies

also had more general limitations, however. First, the systematic review was limited to studies in English only; thus, some studies with adolescents, and potentially other tools, may have been missed. As a result, the generalisability of the findings of this review may be limited as there are some African countries whose official languages are not English. Second, although three geopolitical zones in Nigeria were visited, the participants were drawn mainly from the urban areas where English is usually spoken. Nigeria is a diverse country with over 200 ethnic groups and 500 indigenous languages, therefore, focusing on adapting the screening tools in English restricts its use within Nigeria and outside of non-English speaking African countries. Therefore, the generalisation of the study results may not be ideal. However, English, the official Nigerian language, spoken in most urban cities was chosen for the validation of the screening tools. Replication of these studies in different parts of Nigeria is recommended.

The third is about the issues around ADOS-2, given that there was no focus group review of the ADOS-2 forms. For instance, Abubakar et al. (2014) argues that there is a cultural inappropriateness in some of the stimuli, such as 'TV', 'spoon', and 'sink', and these could have been addressed by a focus group. Another argument presented is around examples used such as 'traffic lights' and behaviours, specifically eye contact. Again, more culturally appropriate, and familiar examples can be used. To resolve these ADOS-2 issues, a focus group discussion as done with the SCQ and SCIL for adaptation purposes is recommended. Fourth, the developmental history of the participants was not assessed as such, so some key information may have been omitted in the classification process, especially for those with intellectual disabilities. Fifth, the responses on the SCQ, unlike the VABS-3, may have been subjective considering the religious and social models of disability which Nigerians exemplify. The VABS-3 is interview-based, i.e., the researcher asks the parent or caregiver questions which afford the opportunity to probe further, unlike the SCQ, which is read by the parent or caregiver, and a yes/no response is selected. This may have impacted the scores on the SCQ. A recommendation will be to administer the SCQ in a quasi-interview to all participants to maintain uniformity (Awadu, 2021).

The test-retest reliability of the SCIL was not examined. Test-retest reliability is one way to examine the stability of tools or individual responses over a time period (Mokinkk et al., 2018a; Weir, 2005; Skirrow et al., 2022). A further assessment of this measurement property within the

Nigerian context should be examined. Despite this limitation, overall, the SCIL and the SCQ had good psychometric properties. The SCQ was utilised only in the urban cities in Nigeria; therefore, the results may not be generalisable to rural areas. There was no dimension reduction of the SCQ through principal component analysis; therefore, the relevance of all the items in the Nigerian context was not examined. However, a CFA was deemed sufficient because the SCQ is well-established and widely used in research. The study results provide evidence to support the retention of the original four factors of the SCQ. Further analysis of the SCQ using another method such as the IRT in the Nigerian or other English-speaking African country is recommended.

Another limitation is the floor effect of the WISC-V and WAIS-III scores. Hessl et al. (2009) noted that intelligence testing in persons with intellectual disabilities was inherently problematic as the normative samples usually did not include an adequate number of participants from the intellectual disabilities' population; thus, the high subjectivity to floor effects. Based on Hessl et al. (2009), 37 participants who could not respond to the WISC-V or WAIS-III due to their level of general intellectual functioning were excluded from the FSIQ analysis. An alternative method of assessing their level of general intellectual functioning, such as calculating normalised scores representing each participant's actual deviation from the standardisation sample using a z-score transformation (Hessl et al., 2009), would have been useful. This was not done. Consequently, we have limited knowledge about how well the screening tools might have worked with the participants with severe to profound intellectual disability who were included in this research.

A final limitation and potential source of conflict is the utilisation of a modified cut-off for the Weschler scales (lowering the FSIQ score for intellectual disability from ≤70 to ≤52 (in study 3a). Studies and debates about intelligence testing of minority and diverse language groups using tests developed and standardised with predominantly White populations abound (Ford, 2005; Lam, 1993; Wicherts et al., 2010; Dramé, & Ferguson, 2019). Africans are included in the minority and diverse language group in this regard. Earlier studies during the eugenics era had concluded that the average IQ for Africans was even lower (below 80) than that of the Western population (Wicherts et al., 2010; Lynn, 2015). In fact, Lynn (2015, p. 45) stated that "the first attempt to estimate the intelligence of Sub-Saharan Africans was made by Sir Francis Galton on the basis of his own experience of them during his travels in southwest Africa and the accounts

of other travelers". He also stated that Galton "constructed a scale of grades of intelligence in which one grade was equivalent to 10.425 IQ points on the IQ scale" (Lynn, 2015, p. 45). He estimated that Africans were about two grades below the English, giving them a mean IQ of 79. Subsequent studies of the level of general intellectual functioning of general population samples of Sub-Saharan Africans have generally maintained this estimate. Basing the IQ of an entire group on "his own experience" is entirely subjective and lacks empirical evidence. If IQ tests were developed in Nigeria, based on the Nigerian culture, it is likely that those in the West would score poorly. In all probability, there is no such thing as a culture free test.

As the popularity of administering standardised tests increases, so will the continued misapplication and disparity in scores in minority groups. Some of these disparities result from deficit perspectives (Ford, Harris, Tyson & Frazier Trotman, 2001). Deficit perspective refers to when individuals "of color who are culturally different from their white counterparts are viewed as culturally deprived or disadvantaged" (Ford et al., 2001, p.52). Revisions of these tests, which include the Weschler scales, have not included sufficient samples from the minority groups to determine the appropriateness of the test items. Whereas there are guidelines for administering these tests, clarity on how to administer them to culturally diverse and minority groups is lacking (Lam, 1993). Suggestions have been made (Lam, 1993; Ford, 2005) on how to mitigate the disparities in scores; some of these are:

a) creating a diverse range of standardised tests that accommodate language and cultural differences. This would be an ideal solution, however, developing many different versions of a test is not feasible in resource-strapped economies due to the bigger problem of technical and practical constraints than those of translation.
b) replacing unfamiliar vocabulary and concepts with material more contextualised and relevant to the person's cultural background and education. For instance, in the WAIS-III, 'Hamlet' could be replaced with 'Things Fall Apart', 'MLK' with 'Fela Anikulapo Kuti', 'Cleopatra' with 'Moremi' and 'prize' with 'gift'.
c) when interpreting test scores, the examiner must consider that many of the existing and traditional tests have not been normed adequately with various cultural groups and be aware of the limitations of standardised tests (see Ford (2005) and Lam (1993) for further discussion of this point).

Tests can be biased if they treat minority and diverse language groups unfairly or discriminate against them, as demonstrated by Galton subjectively assigning an IQ of 79 to Africans. The continued use of the score as a benchmark for Africans indirectly ratifies Galton's position as well as diminishes the potential that exists in the African population. The efforts made by scholars like Wichert and his colleagues to re-examine the low mean IQ assigned to Africans are a step in the right direction. Therefore, the adjustment of the FSIQ score for intellectual disability from ≤70 to ≤52 in this study was based on Wichert et al. (2010) and the suggestions by Lam (1993) and Ford (2005). While we recognise that the adjustment could potentially invalidate norms, which in themselves may be already skewed, test validity is enhanced by considering diversity. Selecting, interpreting and using standardised tests in different environments is complicated, especially when cultural diversity is added on. However, pending further research and the inclusion of more diverse samples in test developments, examiners and professionals should perhaps be more open-minded and amenable to adjustments thereby becoming more inclusive.

Despite these limitations, this research explored and addressed obvious gaps in literature and research, especially in the African context. No previous validation studies had been conducted for autism and intellectual disability screening tools in Nigeria. The validation of the SCQ and SCIL provides the much-needed foundation for screening persons with autism or intellectual disability.

8.8 Conclusions and Recommendations for Future Research

The research aim of this thesis was to identify and validate short screening tools for autism and intellectual disability for use with Nigerian adolescents. In Chapter 1, the theories and models of disability and their impact on research and tool developments were described. The link between the models of disability and the perception of autism or intellectual disability in the Nigerian context was discussed. In Chapter 2, the impact of screening and diagnosing in the Nigerian context based on the models mentioned in Chapter 1 were highlighted, along with the need to identify available screening tools. The reasons for the delayed identification of persons with autism or intellectual disability were also discussed. The findings of the systematic review in Study 1 (Chapter 4) of screening instruments for ID and autism, highlighted the gaps in adolescent studies and the lack of validated tools for use in Nigeria. The results of the systematic

review justified the identification of autism and intellectual disability screening tools for adaptation and use with adolescents within the Nigerian environment. Based on the results of Study 2 (Chapter 5) of the research – the focus group, no material change was made to the SCQ or SCIL, but more relevant examples were included where necessary. The focus group discussions enabled the members to identify problem areas with the screening tools, thereby eliminating problems which may have been encountered during fieldwork. The last stage of the research (Chapter 6 & 7), which involved validating the autism screening tool (SCQ) and the intellectual disability screening tool (SCIL), provided significant evidence for the utilisation of the tools in the Nigerian environment.

Based on the findings of this thesis, some recommendations for further research are made. More studies using the SCQ and SCIL should be conducted in Nigeria and other English-speaking African countries to further test the psychometric properties. Data from this study and future studies should contribute to constructing prevalence data for Nigeria. It is best to estimate the prevalence rate from an actual study of total populations, rather than to extrapolate based on the global rates. Other psychometric properties of the SCIL, such as the test-retest reliability and inter-rater reliability, should be evaluated. Calculating normalised scores for the 37 participants excluded from the FSIQ analyses, representing each participant's actual deviation from the standardisation sample using a z-score transformation for the WISC-V and WAIS-III scores, should be examined with the available data. An additional recommendation is that the ADOS-2 forms and stimuli should be examined using a focus group to ascertain cultural appropriateness or suggest relevant modifications. Also, the Autism Diagnostic Interview-Revised (ADI-R) or DISCO (Diagnostic Interview for Social and Communication Disorders) or 3Di (The Developmental, Dimensional and Diagnostic Interview) should be incorporated into future research as an accompaniment to the ADOS-2.

While the low awareness rate about intellectual disability or autism is similar across Africa, at the core is the lack of validated tools and enough professionals to screen, diagnose and provide intervention to people. A recent study by the Global Research on Developmental Disabilities Collaborators (GRDDC) noted the increase in the global rate of children with developmental disabilities, with intellectual disability contributing 52.4% of the total percentage of children with developmental disabilities in Nigeria. The estimated rate of Nigerian children with

developmental disabilities rose from 1.5 million in 1990 to 2.5 million in 2016, a 66.7% increase (Olusanya et al., 2018). With statistics such as this, it becomes imperative to find ways to improve the process of screening and diagnosing autism and intellectual disability in Nigeria, as an early intervention has shown to be essential to positive outcomes in childhood disorders (Majnemer, 1998; Guralnick, 2005; Guralnick, 2017).

Autism and intellectual disability in Nigeria are being diagnosed through subjective clinical opinions alone, without the use of objective screening tools. It is hoped that in the end, and looking forward to the future, these validated and standardised screening tools for intellectual disabilities (SCIL) and autism spectrum disorder (SCQ) can be used within Nigeria and perhaps other African countries. Having and using these screening tools should hopefully assist the government, healthcare providers, clinicians, educationists, and social care agencies in estimating the number of people with these disabilities within their regions and, in turn, help in developing and targeting resources more effectively.

Finally, identifying individuals with autism or intellectual disabilities requires a framework for supporting them and providing services. There is a growing body of qualified behaviour technicians and analysts in the country who are poised to provide intervention for persons with autism and/or intellectual disability. Therefore, collaborative efforts between the government and these professionals are required to care for this population.

References

Abdulla, M. (1983). How adequate is plasma zinc as an indicator of zinc status? *Progress in Clinical and Biological Research, 129*, 171-183.

Abrams, D. A., Lynch, C. J., Cheng, K. M., Phillips, J., Supekar, K., Ryali, S., … & Menon, V. (2013). Underconnectivity between voice-selective cortex and reward circuitry in children with autism. *Proceedings of the National Academy of Sciences, 110*(29), 12060-12065.

Abubakar, A., de Vries, P., & Newton, C. (2014). *Potential Challenges of Importing Assessment Measures from the West*. Presented at the Autism in Africa Meeting, Accra, Ghana.

Abubakar, A., Ssewanyana, D., de Vries, P. J., & Newton, C. R. (2016a). Autism spectrum disorders in sub-Saharan Africa. *The Lancet Psychiatry, 3*(9), 800-802.

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and revised child behavior profile.* Burlington: Vermont University Associates in Psychiatry.

Achenbach, T., & Rescorla, L. (2001). Manual for the ASEBA school-age forms & profiles. Burlington, VT: University of Vermont. *Research Center for Children, Youth, & Families.*

Adeloye, D., David, R. A., Olaogun, A. A., Auta, A., Adesokan, A., Gadanya, M., ... & Iseolorunkanmi, A. (2017). Health workforce and governance: the crisis in Nigeria. *Human resources for health, 15*(1), 1-8.

Aguwa, C., Carrasco, T., Odongo, N., & Riblet, N. (2022). Barriers to Treatment as a Hindrance to Health and Wellbeing of Individuals with Mental Illnesses in Africa: a Systematic Review. *International journal of mental health and addiction*, 1–17. Advance online publication.

Akobeng, A. K. (2007). Understanding diagnostic tests 1: Sensitivity, specificity and predictive values. *Acta Paediatrica, 96*(3), 338-341.

Akshoomoff, N., Pierce, K., & Courchesne, E. (2002). The neurobiological basis of autism from a developmental perspective. *Development and Psychopathology, 14*(3), 613-634.

Aldosari, M., Fombonne, E., Aldhalaan, H., Ouda, M., Elhag, S., Alshammari, H., ... & Alshaban, F. (2019). Validation of the Arabic version of the Social Communication Questionnaire. *Autism, 23*(7), 1655-1662.

Al Maskari, T. S., Melville, C. A., & Willis, D. S. (2018). Systematic review: Cultural adaptation and feasibility of screening for autism in non-English speaking countries. *International Journal of Mental Health Systems, 12*(1), 1-19.

Alexander, R. M., & Reynolds, M. R. (2020). Intelligence and adaptive behavior: A meta-analysis. *School Psychology Review, 49*(2), 85-110.

Alhojailan, M. I. (2012). Thematic analysis: A critical review of its process and evaluation. *West East Journal of Social Sciences, 1*(1), 39-47.

Allen, A. B., Finestone, M., Eloff, I., Sipsma, H., Makin, J., Triplett, K., ... & Forsyth, B. W. (2014). The role of parenting in affecting the behavior and adaptive functioning of young children of HIV-infected mothers in South Africa. *AIDS and Behavior, 18*, 605-616.

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(2), 202-212.

Al Mamun, K. A., Bardhan, S., Ullah, M. A., Anagnostou, E., Brian, J., Akhter, S., & Rabbani, M. G. (2016). Smart autism - A mobile, interactive and integrated framework for screening and confirmation of autism. Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. *Annual Conference, 2016*, 5989–5992.

Amakom, U., & Ezenekwe, U. (2012). Implications of households catastrophic out of pocket (OOP) healthcare spending in Nigeria. *Journal of Research in Economics and International Finance, 1*(5).

Amakom, U. (2012). Public expenditure on education and healthcare in Nigeria: Who benefits and why? *International Journal of Business and Management*, 7(12), 48.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5 (Fifth Ed.).* Arlington, Virginia. American Psychiatric Publishing.

Ammons, R. B., & Ammons, C. H. (1962). The quick test (QT): Provisional manual. *Psychological Reports*, 11(1), 111-161.

Anderson-Chavarria, M. (2021). The autism predicament: Models of autism and their impact on autistic identity. *Disability & Society*, 1-21.

Ani, C. C., & Grantham-McGregor, S. (1998). Family and personal characteristics of aggressive Nigerian boys: Differences from and similarities with Western findings. *Journal of Adolescent Health*, *23*(5), 311-317.

Ashem, B., & Janes, M. D. (1978). Deleterious effects of chronic undernutrition on cognitive abilities. *Journal of Child Psychology and Psychiatry, 19,* 23-31.

Asperger, H. (1944). Die "Autistischen psychopathen" im kindesalter. *Archiv Für Psychiatrie Und Nervenkrankheiten, 117*(1), 76-136.

Asperger, H. (1991). 'Autistic psychopathy' in childhood. In U. Frith (Ed.), *Autism and Asperger Syndrome* (pp. 37-92). Cambridge: Cambridge University Press.

Astington, J. W., & Edward, M. J. (2010). The development of theory of mind in early childhood. *Encyclopedia on Early Childhood Development, 14*, 1-7.

Asunta, P., Viholainen, H., Ahonen, T., & Rintala, P. (2019). Psychometric properties of observational tools for identifying motor difficulties–a systematic review. *BMC Pediatrics*, *19*(1), 1-13.

Atilola, O., Omigbodun, O., Bella-Awusah, T., Lagunju, I., & Igbeneghu, P. (2014). Neurological and intellectual disabilities among adolescents within a custodial institution in South-West Nigeria. *Journal of Psychiatric and Mental Health Nursing, 21(1),* 31-38.

Awadu, J. E. (2021). *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.* (Thesis). Michigan State University.

Babaknejad, N., Sayehmiri, F., Sayehmiri, K., Mohamadkhani, A., & Bahrami, S. (2016). The relationship between zinc levels and autism: A systematic review and meta-analysis. *Iranian Journal of Child Neurology, 10*(4), 1-9.

Bai, D., Yip, B. H. K., Windham, G. C., Sourander, A., Francis, R., Yoffe, R., et al. (2019). Association of genetic and environmental factors with autism in a 5-country cohort. *JAMA Psychiatry*, *76*(10), 1035-1043.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological medicine*, *25*(1), 63-77.

Baio, J. (2014). Prevalence of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, united states, 2010. *Morbidity and Mortality Weekly Report: Surveillance Summaries, Vol. 63*(2), 1-21.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in south thames: The special needs and autism project (SNAP). *The Lancet, 368*(9531), 210-215.

Bakare, C. G. M. (1972). Social-class differences in the performance of Nigerian children on the Draw-a-Man test. In L. J. Cronbach, & P. J. Drenth (Eds.), *Mental Tests and Cultural Adaptation* (pp. 355–363). The Netherlands: Mouton: The Hague.

Bakare M. O., Ubochi V. N., Ebigbo P. O., & Orovwigho, A. O. (2010). Problem and pro-social behavior among Nigerian children with intellectual disability: The implication for developing policy for school based mental health programs. *Italian Journal of Pediatrics, Vol 36(1), p 37.*

Bakare, M. O., Frazier, T. W., Karpur, A., Abubakar, A., Nyongesa, M. K., Mwangi, P. M., … & Shih, A. (2022). Brief report: Validity and reliability of the Nigerian autism screening questionnaire. *Autism*, 13623613221080250.

Bakare, M. O., & Munir, K. M. (2011). Autism spectrum disorders (ASD) in Africa: A perspective. *African Journal of Psychiatry*, *14*(3), 208-210.

Bakare, M. O., Ebigbo, P. O., Agomoh, A. O., & Menkiti, N. C. (2008). Knowledge about childhood autism among health workers (KCAHW) questionnaire: Description, reliability and internal consistency. *Clinical Practice and Epidemiology in Mental Health: CP & EMH*, *4*, 17-17.

Bakare, M. O., Ubochi, V. N., Okoroikpa, I. N., Aguocha, C. M., & Ebigbo, P. O. (2009). Agreement between clinicians' and care givers' assessment of intelligence in Nigerian children with intellectual disability: 'ratio IQ' as a viable option in the absence of standardized 'deviance IQ' tests in sub-Saharan Africa. *Behavioral and Brain Functions*, *5*(1), 39.

Bakken, T. L., Helverschou, S. B., Eilertsen, D. E., Heggelund, T., Myrbakk, E., & Martinsen, H. (2010). Psychiatric disorders in adolescents and adults with autism and intellectual disability: A representative study in one county in Norway. *Research in Developmental Disabilities*, *31*(6), 1669-1677.

Bargiela, S., Steward, R., & Mandy, W. (2016). The experiences of late-diagnosed women with autism spectrum conditions: An investigation of the female autism phenotype. *Journal of Autism and Developmental Disorders, 46*(10), 3281-3294.

Baron-Cohen, S. (2009). Autism: The empathizing–systemizing (E-S) theory. *Annals of the New York Academy of Sciences, 1156*(1), 68-80.

Baron-Cohen, S., Tsompanidis, A., Auyeung, B., Nørgaard-Pedersen, B., Hougaard, D. M., Abdallah, M., Cohne, A., & Pohl, A. (2020). Foetal oestrogens and autism. *Molecular Psychiatry*, *25*(11), 2970-2978.

Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The Autism Spectrum Quotient (AQ)-adolescent version. *Journal of Autism & Developmental Disorders*, 36(3), 343-350.

Baron-Cohen, S. E., Tager-Flusberg, H. E., & Cohen, D. J. (1994). Understanding other minds: Perspectives from autism. Paper presented at the *Most of the Chapters in this Book were Presented in Draft Form at a Workshop in Seattle, Apr 1991*.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248-254.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism & Developmental Disorders*, 31(1), 5-17.

Barrett, D., Kamenov, K., Pearce, E., & Cieza, A. (2022). Global report on health equity for persons with disabilities. World Health Organization.

Bartak, L., & Rutter, M. (1976). Differences between mentally retarded and normally intelligent autistic children. *Journal of Autism and Childhood Schizophrenia, 6*(2), 109-120.

Barton, M. L., Dumont-Mathieu, T., & Fein, D. (2012). Screening young children for autism spectrum disorders in primary practice. *Journal of Autism and Developmental Disorders, 42*(6), 1165-1174.

Bartunek, J. M., & Murninghan, J. K. (1984). The nominal group technique: Expanding the basic procedure and underlying assumptions. *Group & Organization Studies, 9*(3), 417-432.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186-3191.

Becker, M. M., Wagner, M. B., Bosa, C. A., Schmidt, C., Longo, D., Papaleo, C., & Riesgo, R. S. (2012). Translation and validation of Autism Diagnostic Interview-Revised (ADI-R) for autism diagnosis in Brazil. *Arquivos de Neuro-psiquiatria, 70,* 185-190.

Bedford, K. J. A., & Sharkey, A. B. (2014). Local barriers and solutions to improve care-seeking for childhood pneumonia, diarrhoea and malaria in Kenya, Nigeria and Niger: a qualitative study. *PloS one, 9*(6), e100038.

Bello-Mojeed, M., Omigbodun, O., Bakare, M., & Adewuya, A. (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.*

Belmonte, M. K., Cook, E., Anderson, G. M., Rubenstein, J. L., Greenough, W. T., Beckel-Mitchener, A., et al. (2004). Autism as a disorder of neural information processing: Directions for research and targets for therapy. *Molecular Psychiatry, 9*(7), 646-663.

Bertelli, M. O., Hollenweger Haskell, J., Tassé, M. J., Straccia, C., Rondini, E., Bianco, A., … & Salvador-Carulla, L. (2022). Intellectual Disability/Intellectual developmental disorder. In M. O. Bertelli, S. (. Deb, K. Munir, A. Hassiotis & L. Salvador-Carulla (Eds.), *Textbook of Psychiatry for Intellectual Disability and Autism Spectrum Disorder* (pp. 1-49). Cham: Springer International Publishing.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism Screening Questionnaire: Diagnostic validity. *British Journal of Psychiatry, 175*(5), 444-451.

Billstedt, E. (2007). *Children with autism grow up. Use of the DISCO (Diagnostic Interview for Social and COmmunicative disorders) in population-based cohorts*. Inst of Neuroscience and Physiology. Dept of Psychiatry and Neurochemistry.

Blackwell, T. L., & Madere, L. N. (2005). Slosson Intelligence Test-Revised. *Rehabilitation Counseling Bulletin, 48*(3), 183.

Blaurock-Busch, E., & Nwokolo Chijioke, C. (2018). Heavy metals and trace elements in blood, hair and urine of Nigerian children with autistic spectrum disorder. *International Research Journal of Public Health, 2*:13.

Bogart, K. R., Bonnett, A. K., Logan, S. W., & Kallem, C. (2022). Intervening on disability attitudes through disability models and contact in psychology education. *Scholarship of Teaching and Learning in Psychology, 8*(1), 15.

Bölte, S., Holtmann, M., Poustka, F., Scheurich, A., & Schmidt, L. (2007). Gestalt perception and local-global processing in high-functioning autism. *Journal of Autism and Developmental Disorders, 37*(8), 1493-1504.

Booth, T., Murray, A. L., McKenzie, K., Kuenssberg, R., O'Donnell, M., & Burnett, H. (2013). Brief report: An evaluation of the AQ-10 as a brief screening instrument for ASD in adults. *Journal of Autism and Developmental Disorders, 43*(12), 2997-3000.

Borsboom, D. (2005). *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics*. Cambridge: Cambridge University Press.

Bottini, S. (2018). Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis. *Research in Autism Spectrum Disorders, 45*, 9-26.

Bozalek, F. (2013). Autism screening in children: using the Social Communication Questionnaire in a Western Cape population. (Thesis). University of Cape Town, Faculty of Humanities, Department of Psychology.

Braatveit, K. J., Torsheim, T., & Hove, O. (2018). Screening for intellectual disabilities: A validation of the Hayes Ability Screening Index for in-patients with substance use disorder. *Nordic Journal of Psychiatry, 72*(5), 387-392.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Breidbord, J., & Croudace, T. J. (2013). Reliability generalization for Childhood Autism Rating Scale. *Journal of Autism and Developmental Disorders, 43*(12), 2855-2865.

Brennan, R. L. (2010). Generalizability theory and classical test theory. *Applied Measurement in Education, 24*(1), 1-21.

Brereton, A. V., Tonge, B. J., Mackinnon, A. J., & Einfeld, S. L. (2002). Screening young people for autism with the Developmental Behavior Checklist. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(11), 1369-1375.

Brewer, N., Young, R. L., & Lucas, C. A. (2020). Autism screening in early childhood: Discriminating autism from other developmental concerns. *Frontiers in Neurology, 11*, 594381.

Brooks, W. T., & Benson, B. A. (2013). The validity of the Social Communication Questionnaire in adults with intellectual disability. *Research in Autism Spectrum Disorders, 7*(2), 247-255.

Brown, G. T., & Abdulnabi, H. H. (2017). Evaluating the quality of higher education instructor-constructed multiple-choice tests: Impact on student grades. In *Frontiers in Education, 2(24).* Frontiers Media SA.

Bryson, S. E., Bradley, E. A., Thompson, A., & Wainwright, A. (2008). Prevalence of autism among adolescents with intellectual disabilities. *The Canadian Journal of Psychiatry*, *53*(7), 449-459.

Buijsman, R., Begeer, S., & Scheeren, A. M. (2022). 'Autistic person'or 'person with autism'? Person-first language preference in Dutch adults with autism and parents. *Autism*, 13623613221117914.

Campbell, C. A., & Ashmore, R. J. (1995). Test Review: The Slosson Intelligence Test-Revised (SIT-R). *Measurement and Evaluation in Counseling and Development*, 28(2), 116-18.

Canal-Bedia, R., García-Primo, P., Martín-Cilleros, M. V., Santos-Borbujo, J., Guisuraga-Fernández, Z., Herráez-García, L., ... & Posada-de La Paz, M. (2011). Modified checklist for autism in toddlers: cross-cultural adaptation and validation in Spain. *Journal of Autism and Developmental Disorders*, *41*, 1342-1351.

Cannon, J.' O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in autism spectrum disorder: a systematic review of empirical evidence. *Autism Research*, *14*(4), 604-630.

Cantrill, J., Sibbald, B., & Buetow, S. (1996). The Delphi and Nominal Group Techniques in health services research. *International Journal of Pharmacy Practice*, *4*(2), 67-74.

Carr, A., Linehan, C., O'Reilly, G., Walsh, P. N., & McEvoy, J. (2016). *The Handbook of Intellectual Disability and Clinical Psychology Practice*. Routledge.

Carr, A., & O'Reilly, G. (2016). Diagnosis, classification and epidemiology. In *The Handbook of Intellectual Disability and Clinical Psychology Practice* (pp. 3-44). Routledge.

Cederberg, C. D., Gann, L. C., Foley-Nicpon, M., & Sussman, Z. (2018). ASD screening measures for high-ability youth with ASD: Examining the ASSQ and SRS. *Gifted Child Quarterly*, *62*(2), 220-229.

Centers for Disease Control and Prevention. (2018). *Recommendations and guidelines.*https://www.cdc.gov/ncbddd/autism/hcp-recommendations.html

Central Intelligence Agency. (2023). Nigeria - The World Factbook. Retrieved February 7, 2023, from https://www.cia.gov/the-world-factbook/countries/nigeria/summaries

Charman, T., Baird, G., Simonoff, E., Loucas, T., Chandler, S., Meldrum, D., & Pickles, A. (2007). Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *British Journal of Psychiatry, 191(6),* 554-559.

Charman, T., & Gotham, K. (2013). Measurement issues: Screening and diagnostic instruments for autism spectrum disorders - lessons from research and practise. *Child & Adolescent Mental Health, 18*(1), 52-63.

Chesnut, S. R., Wei, T., Barnard-Brak, L., & Richman, D. M. (2017). A meta-analysis of the Social Communication Questionnaire: Screening for autism spectrum disorder. *Autism, 21*(8), 920-928.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences, 16*(4), 231-239.

Chien, I., Lin, C., Chou, Y., & Chou, P. (2011). Prevalence and incidence of autism spectrum disorders among national health insurance enrollees in Taiwan from 1996 to 2005. *Journal of Child Neurology, 26*(7), 830-834.

Chojnicka, I., & Pisula, E. (2017). Adaptation and validation of the ADOS-2, Polish version. *Frontiers in Psychology, 8,* 1916.

Clements, C. C., Zoltowski, A. R., Yankowitz, L. D., Yerys, B. E., Schultz, R. T., & Herrington, J. D. (2018). Evaluation of the social motivation hypothesis of autism: A systematic review and meta-analysis. *JAMA Psychiatry, 75*(8), 797-808.

Cochrane, A., & Holland, W. (1971). Validation of screening procedures. *British Medical Bulletin, 27*(1), 3-8.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Constant, A., Bervoets, J., Hens, K., & Van de Cruys, S. (2020). Precise worlds for certain minds: an ecological perspective on the relational self in autism. *Topoi*, *39*, 611-622.

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., Metzger, L. M., Shoushtari, C. S., Splinter, R., & Reich, W. (2003). Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of autism and developmental disorders*, *33*, 427-433.

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook Jr, 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.

Cortés, M. J., Orejuela, C., Castellvi, G., Folch, A., Rovira, L., Salvador-Carulla, L., Irazabal, M., Munoz, S., Haro, J. M., Vilella, E. & Martinez-Leal, R. (2018). Psychometric properties of Spanish adaptation of the PDD-MRS scale in adults with intellectual developmental disorders: The EVTEA-DI scale. *Journal of Autism & Developmental Disorders, 48*(5), 1566–1578.

Courchesne, E., Townsend, J., & Saitoh, O. (1994). The brain in infantile autism: Posterior fossa structures are abnormal. *Neurology*, *44*(2), 214-214.

Creamer, D. B. (2008). Disability and Christianity. In D. B. Creamer (Ed.), *Disability and Christian theology: Embodied limits and constructive possibilities* (pp. 35-52). An American Academy of Religion Book. New York, Online edn, Oxford University Press. Retrieved from https://doi.org/10.1093/acprof:oso/9780195369151.003.0003

Critchley, H. D., Daly, E. M., Bullmore, E. T., Williams, S. C., Van Amelsvoort, T., Robertson, D. M., et al. (2000). The functional neuroanatomy of social behaviour: Changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain*, *123*(11), 2203-2212.

Crocker, A. F., & Smith, S. N. (2019). Person-first language: Are we practicing what we preach? *Journal of Multidisciplinary Healthcare, 12*, 125-129.

Cuesta-Gómez, J. L., Andrea Manzone, L., & Posada-De-La-Paz, M. (2016). Modified Checklist for Autism in Toddlers cross-cultural adaptation for Argentina. *International Journal of Developmental Disabilities, 62*(2), 117-123.

Davidson, M. J., Alais, D., van Boxtel, J. J., & Tsuchiya, N. (2018). Attention periodically samples competing stimuli during binocular rivalry. *Elife, 7*, e40868.

de Bildt, A., Sytema, S., Kraijer, D., & Minderaa, R. (2005). Prevalence of pervasive developmental disorders in children and adolescents with mental retardation. *Journal of Child Psychology and Psychiatry, 46*(3), 275-286.

de Vries, P. J. (2016). Thinking globally to meet local needs: autism spectrum disorders in Africa and other low-resource environments. *Current Opinion in Neurology, 29*(2), 130-136.

Deb, S., Dhaliwal, A. J., & Roy, M. (2009). The Usefulness of the DBC-ASA as a screening instrument for autism in children with intellectual disabilities: A pilot study. *Journal of Applied Research in Intellectual Disabilities, 22*(5), 498-501.

Delbecq, A. L., & Van de Ven, Andrew H. (1971). A group process model for problem identification and program planning. *The Journal of Applied Behavioral Science, 7*(4), 466-492.

Delbecq, A. L. (1967). The management of decision-making within the firm: Three strategies for three types of decision-making. *Academy of Management Journal, 10*(4), 329-339.

Demetriou, E. A., DeMayo, M. M., & Guastella, A. J. (2019). Executive function in autism spectrum disorder: History, theoretical models, empirical findings, and potential as an endophenotype. *Frontiers in Psychiatry, 10*, 753.

Deno, S. L. (2005). Problem-solving assessment. In *Assessment for Intervention: A Problem-Solving Approach* (pp. 10–40). New York, NY: The Guilford Press. Retrieved from https://rb.gy/l7zkqt.

DeVellis R. F. (2006). Classical test theory. *Medical care*, *44*(11 Suppl 3), S50–S59.

Dietert, R. R., Dietert, J. M., & DeWitt, J. C. (2011). Environmental risk factors for autism. *Emerging Health Threats Journal*, *4*(1), 7111.

DiLalla, D. L., & Rogers, S. J. (1994). Domains of the Childhood Autism Rating Scale: Relevance for diagnosis and treatment. *Journal of Autism and Developmental Disorders*, *24*(2), 115-128.

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 425-438.

Di Nocera, F., Ferlazzo, F., & Borghi, V. (2001). G theory and the reliability of psychophysiological measures: A tutorial. *Psychophysiology*, *38*(5), 796-806.

Djuwantono, T., Aviani, J. K., Permadi, W., Achmad, T. H., & Halim, D. (2020). Risk of neurodevelopmental disorders in children born from different ART treatments: A systematic review and meta-analysis. *Journal of Neurodevelopmental Disorders*, *12*(1), 1-12.

Dramé, C., & Ferguson, C. J. (2019). Measurements of intelligence in sub-Saharan Africa: Perspectives gathered from research in Mali. *Current Psychology*, *38*(1), 110-115.

Duda, M., Daniels, J., & Wall, D. (2016). Clinical evaluation of a novel and Mobile Autism Risk Assessment. *Journal of Autism & Developmental Disorders*, *46*(6), 1953-1961.

Duncan, A., Ruble, L. A., Meinzen-Derr, J., Thomas, C., & Stark, L. J. (2018). Preliminary efficacy of a daily living skills intervention for adolescents with high-functioning autism spectrum disorder. *Autism*, *22*(8), 983-994.

Durkin, M. (2001). Measurement of childhood disabilities in population studies. Paper presented at the *International Seminar on Measurement of Disability, New York.*

Durkin, M. S., Hasan, Z. M., & Hasan, K. Z. (1995). The Ten Questions screen for childhood disabilities: Its uses and limitations in Pakistan. *Journal of Epidemiology and Community Health, 49*(4), 431-436.

Eaton, R. M. (1921). The value of theories. *The Journal of Philosophy, 18*(25), 682-690.

Ecker, C., Marquand, A., Mourao-Miranda, J., Johnston, P., Daly, E. M., Brammer, M. J., Maltezos, S., Murphy, C. M., Robertson, D., Williams, S. C., & Murphy, D. G. M. (2010). Describing the brain in autism in five dimensions--magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 30*(32), 10612-10623.

Ehlers, S., & Gillberg, C. (1993). The epidemiology of Asperger syndrome: A total population study. *Journal of Child Psychology and Psychiatry, 34*(8), 1327-1350.

Ehlers, S., Gillberg, C., & Wing, L. (1999). A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of Autism and Developmental Disorders, 29*(2), 129-141.

Einfeld, S. L., & Tonge, B. J. (1995). The Developmental Behavior Checklist: The development and validation of an instrument to assess behavioral and emotional disturbance in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders, 25*(2), 81-104.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 439-450.

Emerson, E. (2007). Poverty and people with intellectual disabilities. *Mental retardation and developmental disabilities research reviews, 13*(2), 107-113.

Emerson, E. (2012). Deprivation, ethnicity and the prevalence of intellectual and developmental disabilities. *Journal of Epidemiology and Community Health*, *66*(3), 218-224.

Emerson, E. (2013). Commentary: childhood exposure to environmental adversity and the well-being of people with intellectual disabilities. *Journal of Intellectual Disability Research*, *57*(7), 589-600.

Evans, B. (2013). How autism became autism: The radical transformation of a central concept of child development in Britain. *History of the Human Sciences*, *26*(3), 3.

Fahrmeier, E. D. (1975). The Effect of School Attendance on Intellectual Development in Northern Nigeria. *Child Development, 46(1),* 281–285.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (Fourth ed.) London. Sage.

Fink, A., Kosecoff, J., Chassin, M., & Brook, R. H. (1984). Consensus methods: Characteristics and guidelines for use. *American Journal of Public Health, 74*(9), 979-983.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286.

Folstein, S., & Rutter, M. (1977). Infantile autism: A genetic study of 21 twin pairs. *Journal of Child psychology and Psychiatry*, *18*(4), 297-321.

Ford, D. Y. (2005). Intelligence Testing and Cultural Diversity: Pitfalls and Promises. *Newsletter of The National Research Center on the Gifted and Talented (1990-2013)*. Retrieved February 10, 2023, from https://nrcgt.uconn.edu/newsletters/winter052/

Ford, D. Y., Harris III, J. J., Tyson, C. A., & Trotman, M. F. (2001). Beyond deficit thinking: Providing access for gifted African American students. *Roeper Review*, *24*(2), 52-58.

Ford, G., Andrews, R., Booth, A., Dibdin, J., Hardingham, S., & Kelly, T. P. (2008). Screening for learning disability in an adolescent forensic population. *The Journal of Forensic Psychiatry & Psychology, 19*(3), 371-381.

Franz, L., Chambers, N., von Isenburg, M., & de Vries, P. J. (2017). Autism spectrum disorder in sub-Saharan Africa: A comprehensive scoping review. *Autism Research, 10*(5), 723-749.

Fu, J. M., Satterstrom, F. K., Peng, M., Brand, H., Collins, R. L., Dong, S., ... & Talkowski, M. E. (2022). Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nature genetics*, *54*(9), 1320-1331.

Gaffney, G. R., Tsai, L. Y., Kuperman, S., & Minchin, S. (1987). Cerebellar structure in autism. *American Journal of Diseases of Children*, *141*(12), 1330-1332.

García-Primo, P., Hellendoorn, A., Charman, T., Roeyers, H., Dereu, M., Roge, B., ... & Canal-Bedia, R. (2014). Screening for autism spectrum disorders: state of the art in Europe. *European Child & Adolescent Psychiatry*, *23*, 1005-1021.

Ganglmayer, K., Schuwerk, T., Sodian, B., & Paulus, M. (2020). Do children and adults with autism spectrum condition anticipate others' actions as goal-directed? A predictive coding perspective. *Journal of Autism and Developmental Disorders*, *50*, 2077-2089.

Gau, S. S., Lee, C., Lai, M., Chiu, Y., Huang, Y., Kao, J., et al. (2011). Psychometric properties of the Chinese version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders, 5*(2), 809-818.

Geijsen, K., Kop, N., & de Ruiter, C. (2018). Screening for intellectual disability in Dutch police suspects. *Journal of Investigative Psychology and Offender Profiling, 15*(2), 200-214.

Geschwind, D. H. (2011). Genetics of autism spectrum disorders. *Trends in Cognitive Sciences, 15*(9), 409-416.

Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W., Willemsen, M. H., … & Veltman, J. A. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature, 511*(7509), 344.

Gladstone, M., Lancaster, G. A., Umar, E., Nyirenda, M., Kayira, E., van den Broek, N. R., & Smyth, R. L. (2010). The Malawi Developmental Assessment Tool (MDAT): the creation, validation, and reliability of a tool to assess child development in rural African settings. *PLoS Medicine, 7*(5), e1000273.

Gladstone, M., Mallewa, M., Jalloh, A. A., Voskuijl, W., Postels, D., Groce, N., Kerac, M. & Molyneux, E. (2014). Assessment of neurodisability and malnutrition in children in Africa. *Seminars in Pediatric Neurology*, 21(1), 50-57.

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(3), 173-179.

Gopalakrishnan, S., & Ganeshkumar, P. (2013). Systematic reviews and meta-analysis: Understanding the best evidence in primary healthcare. *Journal of Family Medicine and Primary Care, 2*(1), 9-14.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37*, 613-627.

Grabinski, M., & Klinkova, G. (2020). Scrutinizing distributions proves that IQ is inherited and explains the fat tail. *Applied Mathematics*, *11*(10), 957.

Grabrucker, A. M. (2013). Environmental factors in autism. *Frontiers in Psychiatry, 3*, 118.

Grinker, R. R., Kang-Yi, C. D., Ahmann, C., Beidas, R. S., Lagman, A., & Mandell, D. S. (2015). Cultural adaptation and translation of outreach materials on autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(8), 2329-2336.

Grisay, A. (2003). Translation procedures in OECD/PISA 2000 international assessment. *Language Testing, 20*(2), 225-240.

Grzadzinski, R., Huerta, M., & Lord, C. (2013). DSM-5 and autism spectrum disorders (ASDs): An opportunity for identifying ASD subtypes. *Molecular Autism, 4*(1), 1-6.

Guevara, J. P., Gerdes, M., Localio, R., Huang, Y. V., Pinto-Martin, J., Minkovitz, C. S., et al. (2013). Effectiveness of developmental screening in an urban setting. *Pediatrics, 131*(1), 30.

Guion, L. A., Diehl, D. C., & McDonald, D. (2011). Triangulation: Establishing the validity of qualitative studies. *Edis, 2011*(8), 3-3.

Guo, Y., Tang, Y., Rice, C., Lee, L., Wang, Y., & Cubells, J. F. (2011). Validation of the Autism Spectrum Screening Questionnaire, Mandarin Chinese version (CH-ASSQ) in Beijing, China. *Autism, 15*(6), 713-727.

Gura, G. F., Champagne, M. T., & Blood-Siegfried, J. (2011). Autism spectrum disorder screening in primary care. *Journal of Developmental & Behavioral Pediatrics, 32*(1).

Guralnick, M. J. (2005). Early intervention for children with intellectual disabilities: Current knowledge and future prospects. *Journal of Applied Research in Intellectual Disabilities., 18*(4), 313-324.

Guralnick, M. J. (2017). Early intervention for children with intellectual disabilities: An update. *Journal of Applied Research in Intellectual Disabilities, 30*(2), 211-229.

Gureje, O., & Lasebikan, V. O. (2006). Use of mental health services in a developing country. *Social Psychiatry and Psychiatric Epidemiology, 41*(1), 44-49.

Ha, S., Sohn, I., Kim, N., Sim, H. J., & Cheon, K. (2015). Characteristics of brains in autism spectrum disorder: Structure, function and connectivity across the lifespan. *Experimental Neurobiology, 24*(4), 273-284.

Hacker, K. A., Penfold, R., Arsenault, L., Zhang, F., Murphy, M., & Wissow, L. (2014). Screening for behavioral health issues in children enrolled in Massachusetts Medicaid. *Pediatrics, 133*(1), 46-54.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Vol. 2). Sage.

Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational measurement: issues and practice*, *12*(3), 38-47.

Hambleton, R. K. (1996). Guidelines for adapting educational and psychological tests. Paper presented at the *Paper Presented at the Annual Meeting of the National Council on Measurement in Education (New York, NY, April 9-11, 1996).* Retrieved from https://files.eric.ed.gov/fulltext/ED399291.pdf

Hambleton, R. K. (2002). Adapting achievement tests into multiple languages for international assessments. In Porter, A.C. and Gamoran, A. editors. Methodological advances in cross national surveys of educational achievement. (pp. 58-79). Washington DC: National Academy Press Washington.

Happé, F. G. (1994). Annotation: Current psychological theories of autism: The "theory of mind" account and rival theories. *Child Psychology & Psychiatry & Allied Disciplines*, *35(2)*, 215-229.

Happé, F. G. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbol*, *10*(4), 275-295.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5-25.

Hasegawa, J., Ito, Y. M., & Yamauchi, T. (2017). Development of a screening tool to predict malnutrition among children under two years old in Zambia. *Global Health Action*, *10*(1), 1339981.

Hassiotis, A., Strydom, A., Hall, I., Ali, A., Lawrence-Smith, G., Meltzer, H., ... & Bebbington, P. (2008). Psychiatric morbidity and social functioning among adults with borderline intelligence living in private households. *Journal of Intellectual Disability Research*, *52*(2), 95-106.

Hayes, S. C. (2000). *Hayes ability screening index: HASI: Manual*. University of Sydney.

Hayes, S. C. (2002). Early intervention or early incarceration? Using a screening test for intellectual disability in the criminal justice system. *Journal of Applied Research in Intellectual Disabilities, 15*(2), 120-128.

Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., ... & Statistical Analysis Gu Core H. 17. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature*, *542*(7641), 348-351.

Heinrich, M., Böhm, J., & Sappok, T. (2018). Diagnosing autism in adults with intellectual disability: Validation of the DiBAS-R in an independent sample. *Journal of Autism & Developmental Disorders, 48*(2), 341-350.

Hertz-Picciotto, I., Schmidt, R. J., & Krakowiak, P. (2018). Understanding environmental contributions to autism: Causal concepts and the state of science. *Autism Research, 11*(4), 554-586.

Hessl, D., Nguyen, D. V., Green, C., Chavez, A., Tassone, F., Hagerman, R. J., Senturk, D., Schneider, A., Lightbody, A., Reiss, A. L. & Hall, S. (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: The case of fragile X syndrome. *Journal of Neurodevelopmental Disorders, 1*(1), 33-45.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*(2), 189-233.

Hill, E. L., & Frith, U. (2003). Understanding autism: Insights from mind and brain. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*(1430), 281-289.

Hintze, J. M., Owen, S. V., Shapiro, E. S., & Daly III, E. J. (2000). Generalizability of oral reading fluency measures: Application of G theory to curriculum-based measurement. *School Psychology Quarterly*, *15*(1), 52.

Hirota, T., So, R., Kim, Y. S., Leventhal, B., & Epstein, R. A. (2018). A systematic review of screening tools in non-young children and adults for autism spectrum disorder. *Research in Developmental Disabilities, 80*, 1-12.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

Huang, J., Zhu, T., Qu, Y., & Mu, D. (2016). Prenatal, perinatal and neonatal risk factors for intellectual disability: a systemic review and meta-analysis. *PloS One, 11*(4), e0153655.

Humphrey-Murto, S., Varpio, L., Gonsalves, C., & Wood, T. J. (2017). Using consensus group methods such as Delphi and nominal group in medical education research. *Medical Teacher, 39*(1), 14-19.

Hutchings, H. A., Rapport, F. L., Wright, S., Doel, M. A., & Wainwright, P. (2010). Obtaining consensus regarding patient-centred professionalism in community pharmacy: Nominal group work activity with professionals, stakeholders and members of the public. *International Journal of Pharmacy Practice, 18*(3), 149-158.

Hutchings, H., Rapport, F., Wright, S., Doel, M., & Jones, A. (2012). Obtaining consensus about patient-centred professionalism in community nursing: Nominal group work activity with professionals and the public. *Journal of Advanced Nursing, 68*(11), 2429-2442.

Idring, S., Lundberg, M., Sturm, H., Dalman, C., Gumpert, C., Rai, D., Lee, B. K., & Magnusson, C. (2015). Changes in prevalence of autism spectrum disorders in 2001–2011: Findings from the Stockholm youth cohort. *Journal of Autism and Developmental Disorders, 45*(6), 1766-1773.

Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American sociological review*, 19-51.

International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests (Second Edition),* Retrieved from www.InTestCom.org

Iragorri, N., & Spackman, E. (2018). Assessing the value of screening tools: Reviewing the challenges and opportunities of cost-effectiveness analysis. *Public Health Reviews*, *39*(1), 17.

Iwase, S., Bérubé, N. G., Zhou, Z., Kasri, N. N., Battaglioli, E., Scandaglia, M., & Barco, A. (2017). Epigenetic etiology of intellectual disability. *Journal of Neuroscience*, *37*(45), 10773-10782.

Jabrayilov, R., Emons, W. H. M., & Sijtsma, K. (2016). Comparison of Classical Test Theory and Item Response Theory in Individual Change Assessment. *Applied psychological measurement*, *40*(8), 559–572.

Jadhav, D. (2009). What is cultural validity and why is it ignored?. AMB Diemen.

Jinabhai, C. C., Taylor, M., Rangongo, M. F., Mkhize, N. J., Anderson, S., Pillay, B. J., & Sullivan, K. R. (2004). Investigating the mental abilities of rural Zulu primary school children in South Africa. *Ethnicity & health*, 9(1), 17-36.

Jones, J., & Hunter, D. (1995). Consensus methods for medical and health services research. *BMJ (Clinical Research Ed.)*, *311*(7001), 376-380.

Joseph, R. M. (1999). Neuropsychological frameworks for understanding autism. *International Review of Psychiatry*, *11*(4), 309-324.

Joseph, R. M., & Tager–Flusberg, H. (2004). The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and Psychopathology*, *16*(1), 137-155.

Kaal, H.L., Nijman, H.L.I. and Moonen, X.M.H. (2015), "Identifying offenders with an intellectual disability in detention in The Netherlands". *Journal of Intellectual Disabilities and Offending Behaviour*, Vol. 6 No. 2, pp. 94-101.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36.

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.

Kamau, L. Z. (2017). Autism spectrum disorders (ASD) in Kenya: Barriers encountered in diagnosis, treatment and management. *Journal of Research in Pharmaceutical Science, 3,* 1-11.

Kamin, L. J. (2006). African IQ and mental retardation. *South African Journal of Psychology, 36*(1), 1-9.

Kanne, S. M., Gerber, A. J., Quirmbach, L. M., Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2011). The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *Journal of Autism and Developmental Disorders, 41*(8), 1007-1018.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217-250.

Kapungwe, A., Cooper, S., Mwanza, J., Mwape, L., Sikwese, A., Kakuma, R., Lund, C., Flisher, A. J. & MHaPP Research Programme Consortium (2010). Mental illness-stigma and discrimination in Zambia. *African Journal of Psychiatry, 13*(3).

Karaminis, T., & Stavrakaki, S. (2022). The psychometric properties of the Greek version of the Social Communication Questionnaire. *Autism Research, 15*(9), 1768-1780.

Karvelis, P., Seitz, A. R., Lawrie, S. M., & Seriès, P. (2018). Autistic traits, but not schizotypy, predict increased weighting of sensory information in Bayesian visual integration. *ELife, 7,* e34115.

Kayser-Jones, J., Abu-Saad, H., & Akinnaso, F. N. (1982). Nigeria: The land, its people, and health care. *Journal of Nursing Education, 21*(7), 32-7.

Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., ... & Rahman, A. (2011). Child and adolescent mental health worldwide: evidence for action. *The Lancet*, *378*(9801), 1515-1525.

Kim, J. H., Sunwoo, H. J., Park, S. B., Noh, D. H., Jung, Y. K., Cho, I. H., ... & Yoo, H. J. (2015). A validation study of the Korean version of Social Communication Questionnaire. *Journal of the Korean Academy of Child and Adolescent Psychiatry*, *26*(3), 197-208.

Kishore, M. T., & Seshadri, S. P. (2019). Clinical practice guidelines for assessment and management of intellectual disability. *Indian Journal of Psychiatry, 61*(Suppl 2), 194-210.

Kleinman, J. M., Ventola, P. E., Pandey, J., Verbalis, A. D., Barton, M., Hodgson, S., Green, J., Dumont-Mathieu, T., Robins, D. L., & Fein, D. (2008). Diagnostic stability in very young children with autism spectrum disorders. *Journal of autism and developmental disorders*, *38*(4), 606–615.

Knox, J., Arpadi, S. M., Kauchali, S., Craib, M., Kvalsvig, J. D., Taylor, M., et al. (2018). Screening for developmental disabilities in HIV positive and HIV negative children in south Africa: Results from the Asenze study. *PloS One*, *13*(7), e0199860.

Koegel, L. K., Koegel, R. L., Ashbaugh, K., & Bradshaw, J. (2014). The importance of early identification and intervention for children with or at risk for autism spectrum disorders. *International Journal of Speech-Language Pathology*, *16*(1), 50-56.

Koğar, H., & Koğar, E. Y. (2015). Comparison of different estimation methods for categorical and ordinal data in confirmatory factor analysis. *Journal of Measurement and Evaluation in Education and Psychology, 6*(2).

Kopp, S., & Gillberg, C. (2011). The Autism Spectrum Screening Questionnaire (ASSQ)-Revised Extended Version (ASSQ-REV): An instrument for better capturing the autism

phenotype in girls? A preliminary study involving 191 clinical cases and community controls. *Research in Developmental Disabilities, 32(6),* 2875–2888.

Koura, K. G., Boivin, M. J., Davidson, L. L., Ouédraogo, S., Zoumenou, R., Alao, M. J., Garcia, A., Massougbodji, A., Cot, M., & Bodeau-Livinec, F. (2013). Usefulness of child development assessments for low-resource settings in francophone Africa. *Journal of Developmental and Behavioral Pediatrics: JDBP, 34*(7), 486-493.

Kraijer, D. W. (1990). AVZ, Autisme-en Verwante kontaktstoornissenschaal voor Zwakzinnigen. Handleiding (Autism and related contact disorders scale for the mentally retarded. Manual). Swets and Zeitlinger, Lisse, The Netherlands.

Kraijer, D. W. (1994). AVZ-R, Autisme-en Verwante stoornissenschaal voor Zwakzinnigen-Revisie. Handleiding. Sterk herziene en uitgebreide uitgave (Autism and related disorders scale for the mentally retarded-Revision. Manual. Swets and Zeitlinger Lisse, The Netherlands.

Kraijer, D., & de Bildt, A. (2005). The PDD-MRS: An instrument for identification of autism spectrum disorders in persons with mental retardation. *Journal of Autism and Developmental Disorders, 35*(4), 499-513.

Kromberg, J., Zwane, E., Manga, P., Venter, A., Rosen, E., & Christianson, A. (2008). Intellectual disability in the context of a South African population. *Journal of policy and practice in intellectual disabilities, 5(2),* 89-95.

Krug, D. A., Arick, J., & Almond, P. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry, 21*(3), 221-229.

Küçükdeveci, A. A., Sahin, H., Ataman, S., Griffiths, B., & Tennant, A. (2004). Issues in cross-cultural validity: Example from the adaptation, reliability, and validity testing of a Turkish version of the Stanford Health Assessment Questionnaire. *Arthritis Care & Research, 51*(1), 14-19.

Kunnan, A. J. (1992). An investigation of a criterion-referenced test using G-theory, and factor and cluster analyses. *Language Testing*, *9*(1), 30-49.

Kunen, S., Overstreet, S., & Salles, C. (1996). Concurrent validity study of the Slosson Intelligence Test-Revised in mental retardation testing. *Mental Retardation, 34*(6), 380-386.

La Malfa, G., Lassi, S., Bertelli, M., Salvini, R., & Placidi, G. (2004). Autism and intellectual disability: A study of prevalence on a sample of the Italian population. *Journal of Intellectual Disability Research, 48*(3), 262-267.

Lam, T. C. M. (1993). Testability: A critical issue in testing language minority students with standardized achievement. *Measurement & Evaluation in Counseling & Development, 26*(3), 179-191.

LaSalle, J. M., Vallero, R. O., & Mitchell, M. M. (2013). Epigenetics at the interface of genetics and environmental factors in autism. *Environmental Epigenomics in Health and Disease*, 97-114.

Lasko, T. A., Bhagwat, J. G., Zou, K. H., & Ohno-Machado, L. (2005). The use of receiver operating characteristic curves in biomedical informatics. *Journal of Biomedical Informatics, 38*(5), 404-415.

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience, 8*, 302.

Lebersfeld, J. B., Swanson, M., Clesi, C. D., & O'Kelley, S. E. (2021). Systematic review and meta-analysis of the clinical utility of the ADOS-2 and the ADI-R in diagnosing autism spectrum disorders in children. *Journal of Autism and Developmental Disorders, 51*(11), 4101-4114.

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism and Developmental Disorders, 37*, 894-910.

Lesinskiene, S. (2000). Vilniaus miesto vaiku autizmas, Vilniaus Universitetas, Daktaro disetacijos santrauka, Biomedicinos mokslai, medicina 07B, Psichiatrija B650. Vilnius, Lithuania.

Leslie, A. M., & Happé, F. (1989). Autism and ostensive communication: The relevance of metarepresentation. *Development and Psychopathology, 1*(3), 205-212.

Levy, F. (2007). Theories of autism. *Australian & New Zealand Journal of Psychiatry, 41*(11), 859-868.

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological Assessment* Oxford University Press, USA.

Limbos, M. M., & Joyce, D. P. (2011). Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *Journal of Developmental & Behavioral Pediatrics, 32*(7).

List, D. (2001). The consensus group technique in social research. *Field methods*, *13*(3), 277-290.

Liu, S., Wang, X., Chen, Q., Chen, J., Jin, C., Zhan, X., Guo, C., Li, X., Lin, L., & Jing, J. (2022). The validity and reliability of the simplified Chinese version of the Social Communication Questionnaire. *Autism Research*, 1-10.

Livingston, L. A., Shah, P., & Happé, F. (2019). Compensation in autism is not consistent with social motivation theory. *Behavioral and Brain Sciences, 42*, e99.

Loannou, C., Seernani, D., Stefanou, M. E., Biscaldi-Schaefer, M., Tebartz Van Elst, L., Fleischhaker, C., ... & Klein, C. (2020). Social visual perception under the eye of bayesian theories in autism spectrum disorder using advanced modeling of spatial and temporal parameters. *Frontiers in Psychiatry, 11*, 585149.

Long, K. A., Gordillo, M., & Orsmond, G. I. (2020). Improving the validity and generalizability of adult autism research through incorporating family and cultural contexts. *Autism in Adulthood, 2*(3), 177-184.

Lopez, B. R., Lincoln, A. J., Ozonoff, S., & Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of autistic disorder. *Journal of Autism and Developmental Disorders, 35*(4), 445-460.

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). ADOS-2: Autism Diagnostic Observation Schedule, (ADOS-2). *Torrance, CA: Western Psychological Services.*

Lord, C., & Jones, R. M. (2012). Annual research review: Re-thinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 53*(5), 490-509.

Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D. M., ... & Risi, S. (2012). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of general psychiatry*, *69*(3), 306-313.

Lotter, V. (1978). Childhood autism in Africa. *Journal of Child Psychology and Psychiatry, 19*(3), 231-244.

Luckasson, R., & Schalock, R. L. (2013). What's at stake in the lives of people with intellectual disability? part II: Recommendations for naming, defining, diagnosing, classifying, and planning supports. *Intellectual and Developmental Disabilities, 51*(2), 94-101.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., Bishop, S., Esler, A., Hus, V., Oti, R., Richler, J., Risi, S., & Lord, C. (2009). The Autism Diagnostic Observation Schedule-toddler module: a new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of autism and developmental disorders*, *39*(9), 1305–1320.

Lynn, R. (2015). *Race differences in intelligence: An evolutionary analysis.* (Second Ed.) Washington Summit Publishers. Retrieved from https://www.intelligence-humaine.com/wp-

content/uploads/2019/03/Race-Differences-in-Intelligence-second-edition-2015-1.pdf Accessed 21 November 2022.

Maenner, M., Shaw, K., Bakian, A., Bilder, D., Durkin, M., Esler, A., … Cogswell, M. E. (2018). Prevalence and characteristics of autism spectrum disorder among children aged 8 Years—Autism and developmental disabilities monitoring network, 11 sites. *United States, 70*, 11-20.

Majnemer, A. (1998). Benefits of early intervention for children with developmental disabilities. In *Seminars in Pediatric Neurology* (Vol. 5, No. 1, pp. 62-69). WB Saunders.

Malcolm-Smith, S., Hoogenhout, M., Ing, N., Thomas, K. G., & de Vries, P. (2013). Autism spectrum disorders—Global challenges and local opportunities. *Journal of Child & Adolescent Mental Health*, *25*(1), 1-5.

Mallett, R., Hagen-Zanker, J., Slater, R., & Duvendack, M. (2012). The benefits and challenges of using systematic reviews in international development research. *Journal of Development Effectiveness*, *4*(3), 445-455.

Marks, D. (1997). Models of disability. *Disability and Rehabilitation*, *19*(3), 85-91.

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. *Autism Research*, *12*(2), 176-199.

Marotta, R., Risoleo, M. C., Messina, G., Parisi, L., Carotenuto, M., Vetri, L., & Roccella, M. (2020). The neurochemistry of autism. *Brain sciences, 10(3),* 163.

Mattila, M. L., Jussila, K., Kuusikko, S., Kielinen, M., Linna, S. L., Ebeling, H., ... & Moilanen, I. (2009). When does the Autism Spectrum Screening Questionnaire (ASSQ) predict autism spectrum disorders in primary school-aged children?. *European Child & Adolescent Psychiatry*, *18*, 499-509.

Matson, J. L., & Goldin, R. L. (2013). Comorbidity and autism: Trends, topics and future directions. *Research in Autism Spectrum Disorders, 7(10),* 1228-1233.

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in developmental disabilities, 30(6),* 1107-1114.

Matsumoto, D., & Yoo, S. H. (2006). Toward a new generation of cross-cultural research. *Perspectives on Psychological Science, 1*(3), 234-250.

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: A meta-analysis of population-based studies. *Research in Developmental Disabilities, 32*(2), 419-436.

Maxim, L. D., Niebo, R., & Utell, M. J. (2014). Screening tests: A review with examples. *Inhalation Toxicology, 26*(13), 811-828.

Mazefsky, C. A., Anderson, R., Conner, C. M., & Minshew, N. (2011). Child Behavior Checklist scores for school-aged children with autism: Preliminary evidence of patterns suggesting the need for referral. *Journal of Psychopathology and Behavioral Assessment, 33*(1), 31-37.

McBride, O., Heslop, P., Glover, G., Taggart, T., Hanna-Trainor, L., Shevlin, M., & Murphy, J. (2021). Prevalence estimation of intellectual disability using national administrative and household survey data: the importance of survey question specificity. *International Journal of Population Data Science, 6*(1).

McConachie, H., Mason, D., Parr, J. R., Garland, D., Wilson, C., & Rodgers, J. (2018). Enhancing the validity of a quality of life measure for autistic people. *Journal of Autism and Developmental Disorders, 48*(5), 1596-1611.

McConachie, H., Parr, J. R., Glod, M., Hanratty, J., Livingstone, N., Oono, I. P., ... & Williams, K. (2015). Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technology Assessment, 19*(41), 1-538.

McGlashan, T. H. (2011). Eugen Bleuler: Centennial anniversary of his 1911 publication of Dementia Praecox or the group of schizophrenias. *Schizophrenia bulletin*, *37*(6), 1101-1103.

McKenzie, K., & Murray, A. L. (2015). Evaluating the use of the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q) to estimate IQ in children with low intellectual ability. *Research in Developmental Disabilities, 37*, 31-36.

McKenzie, K. and Paxton, D. (2006a). *Learning Disability Screening Questionnaire*. GCM Records, Edinburgh.

McKenzie, K., & Paxton, D. (2006b). Promoting access to services: The development of a new screening tool. *Learning Disability Practice, 9(6),* 17–21.

McKenzie, K., Paxton, D., Michie, A., Murray, G., Murray, A., & Curtis, J. (2012a). Screening with young offenders with an intellectual disability. *Journal of Forensic Psychiatry & Psychology*, *23*(5-6), 676-688.

McKenzie, K., Michie, A., Murray, A., & Hales, C. (2012b). Screening for offenders with an intellectual disability: the validity of the Learning Disability Screening Questionnaire. *Research in Developmental Disabilities, 33(3),* 791-795.

McKenzie, K., Murray, A. L., Murray, K. R., & Murray, G. C. (2014). Assessing the accuracy of the WISC-IV seven-subtest short form and the child and adolescent intellectual disability screening questionnaire in identifying intellectual disability in children. *Child Neuropsychology, 20(3),* 372–377.

McKenzie, K., Murray, G., Murray, A., Delahunty, L., Hutton, L., Murray, K., & O'Hare, A. (2019). Child and Adolescent Intellectual Disability Screening Questionnaire to identify children with intellectual disability. *Developmental Medicine & Child Neurology*, 61(4), 444-450

McKenzie, K., Paxton, D., Murray, G., Milanesi, P., & Murray, A. L. (2012c). The evaluation of a screening tool for children with an intellectual disability: The Child and Adolescent

Intellectual Disability Screening Questionnaire. *Research in Developmental Disabilities, 33*(4), 1068-1075.

McKenzie, K. and Paxton D. (2012). *Child and Adolescent Intellectual Disability Screening Questionnaire*. GCM Records, Edinburgh.

McKenzie, K., Milton, M., Smith, G., & Ouellette-Kuntz, H. (2016). Systematic review of the prevalence and incidence of intellectual disabilities: Current trends and issues. *Current Developmental Disorders Reports, 3*(2), 104-115.

McMillan, S. S., Kelly, F., Sav, A., Kendall, E., King, M. A., Whitty, J. A., & Wheeler, A. J. (2014). Using the nominal group technique: How to analyse across multiple groups. *Health Services and Outcomes Research Methodology, 14*(3), 92-108.

McMillan, S. S., King, M., & Tully, M. P. (2016). How to use the Nominal Group and Delphi techniques. *International Journal of Clinical Pharmacy, 38*(3), 655-662.

Merriam-Webster. (n.d.). *Disability. In Merriam-Webster.com dictionary*. Retrieved December 27, 2022 https://www.merriam-webster.com/dictionary/disability

Merritt, C. J. (2012). The empathizing–systemizing (E–S) model of autism and psychoanalytic theories of truth, play and symbolization. *Psychoanalytic Psychotherapy, 26*(4), 327-337.

Mesibov, G. B., Schopler, E., Schaffer, B., & Michal, N. (1989). Use of the Childhood Autism Rating Scale with autistic adolescents and adults. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*(4), 538-541.

Mesulam, M. (2002). The human frontal lobes: Transcending the default mode through contingent encoding. *Principles of Frontal Lobe Function, 54*, 8-30.

Meyer, A. T., Powell, P. S., Butera, N., Klinger, M. R., & Klinger, L. G. (2018). Brief report: Developmental trajectories of adaptive behavior in children and adolescents with ASD. *Journal of Autism and Developmental Disorders, 48*(8), 2870-2878.

Milton, D. (2012). So what exactly is autism? Retrieved from https://www.researchgate.net/profile/Dinah-Murray-2/publication/338127881_Monotropism_-_An_Interest_Based_Account_of_Autism/links/5e49da7b458515072da455f4/Monotropism-An-Interest-Based-Account-of-Autism.pdf

Mîndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society, 1*(1), 60-66.

Minshew, N. J., Johnson, C., & Luna, B. (2000). The cognitive and neural basis of autism: A disorder of complex information processing and dysfunction of neocortical systems. *International review of research in mental retardation* (pp. 111-138) Elsevier.

Mirabi M. (1985). Psychiatry and mental retardation: A historical perspective. In Pichot P., Berner P., Wolf R. & Thau K. (Eds.), *Child and Adolescent Psychiatry, Mental Retardation, and Geriatric Psychiatry*. (pp. 171-181). Boston, MA: Springer.

Mohamed, F. E., Zaky, E. A., Youssef, A., Elhossiny, R., Zahra, S., Khalaf, R., ... & Eldin, W. S. (2016). Screening of Egyptian toddlers for autism spectrum disorder using an Arabic validated version of M-CHAT; report of a community-based study (Stage I). *European Psychiatry*, *34*, 43-48.

Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2018a). COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). *User manual. https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf* Accessed 28 November 2018.

Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018b). COSMIN risk of Bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1171–1179.

Molnár, Z. (May 2004). Thomas Willis (1621–1675), the founder of clinical neuroscience. *Nature Reviews Neuroscience, 5*(4), 329-335.

Morales-Hidalgo, P., Hernández-Martínez, C., Voltas, N., & Canals, J. (2017). EDUTEA: A DSM-5 teacher screening questionnaire for autism spectrum disorder and social pragmatic communication disorder. *International Journal of Clinical Health & Psychology, 17*(3), 269-281.

Moran, J. M. (2013). Lifespan development: The effects of typical aging on theory of mind. *Behavioural Brain Research, 237*, 32-40.

Moskowitz, A., & Heim, G. (2011). Eugen Bleuler's dementia praecox or the group of schizophrenias (1911): A centenary appreciation and reconsideration. *Schizophrenia Bulletin, 37*(3), 471-479.

Mouti, A., Dryer, R., & Kohn, M. (2019). Differentiating autism spectrum disorder from ADHD using the Social Communication Questionnaire. *Journal of Attention Disorders*, *23*(8), 828-837.

Mung'ala-Odera, V., Meehan, R., Njuguna, P., Mturi, N., Alcock, K., Carter, J. A., & Newton, C. R. (2004). Validity and reliability of the 'ten questions' questionnaire for detecting moderate to severe neurological impairment in children aged 6-9 years in rural Kenya. *Neuroepidemiology*, 23(1-2), 67-72.

Murphy, M., Black, N., Lamping, D., McKee, C., Sanderson, C., Askham, J., & Marteau, T. (1998). Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment (Winchester, England), 2*(3), i-88.

Murphy, G. H., Gardner, J., & Freeman, M. J. (2017). Screening prisoners for intellectual disabilities in three English prisons. *Journal of Applied Research in Intellectual Disabilities 30(1),* 198-204.

Murray, D. (2018). Monotropism–an interest based account of autism. *Encyclopedia of Autism Spectrum Disorders, 10*, 971-978.

Murray, D., Lesser, M., & Lawson, W. (2005). Attention, monotropism and the diagnostic criteria for autism. *Autism, 9*(2), 139-156.

Musser, G. (2019). The predictive coding theory of autism, explained. *Spectrum | Autism Research News, 1-3*. Retrieved from https://www.spectrumnews.org/news/predictive-coding-theory-autism-explained/#:~:text=The%20predictive%20coding%20theory%20of%20autism%20proposes%20that%20an%20autistic,unable%20to%20tune%20it%20out.

Nah, Y., Young, R. L., Brewer, N., & Berlingeri, G. (2014). Autism Detection in Early Childhood (ADEC): Reliability and validity data for a level 2 screening tool for autistic disorder. *Psychological Assessment, 26*(1), 215.

Narby, J. (1982). The evolution of attitudes towards mental illness in pre-industrial England. *Orthomolecular.Org, 11*(2), February 28, 2019-pp 103-110.

Nash, C., Hawkins, A., Kawchuk, J., & Shea, S. E. (2012). What's in a name? Attitudes surrounding the use of the term 'mental retardation'. *Paediatrics & Child Health, 17*(2), 71-74.

National Institute for Health and Care Excellence. (2017). "*Autism spectrum disorder in under 19s: recognition, referral and diagnosis.*" (CG128). ttps://www.nice.org.uk/guidance/cg128 Accessed 11 February 2023.

Nenty, H. J., & Dinero, T. E. (1981). A cross-cultural analysis of the fairness of the Cattell Culture Fair Intelligence Test using the Rasch model. *Applied Psychological Measurement, 5*(3), 355-368.

Newsom, J. (2018a). Alternative estimation methods (psy 523/623 structural equation modeling, spring 2018). *Manuscript Retrieved from Http://web.Pdx.edu/~newsomj/semclass/*,

Newsom, J. (2018b). Some clarifications and recommendations on fit indices (psy 523/623 structural equation modeling, spring 2018). *Manuscript Retrieved from Http://web.Pdx.edu/~newsomj/semclass/*,

Nijman, H., Kaal, H., van Scheppingen, L., & Moonen, X. (2018). Development and testing of a Screener for Intelligence and Learning Disabilities (SCIL). *Journal of Applied Research in Intellectual Disabilities, 31*(1), e59-e67.

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.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1609406917733847.

Nwanze, H. O., & Okeowo, P. A. (1980). The usefulness of a developmental profile in predicting reading retardation in Nigerian children. *West African Journal of Educational & Vocational Measurement, 5(1),* 50–59.

Nwokolo, E. (2017). *Barriers to assessment and intervention for people with autism in Nigeria.* Unpublished MSc, University of Kent, Tizard Centre, Canterbury.

Nwokolo, E. U., Langdon, P. E., & Murphy, G. H. (2022). Screening for intellectual disabilities and/or autism amongst older children and young adults: A systematic review of tools for use in Africa. *Review Journal of Autism and Developmental Disorders*. Advance online publication.

Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E., (in press). Using the consensus group method to select the best screening tools for autism and intellectual disability for use with Nigerian adolescents. *Journal of Policy & Practice in Intellectual Disabilities.*

O'Donnell, O. (2007). Access to health care in developing countries: Breaking down demand side barriers. *Cadernos De Saude Publica, 23*, 2820-2834.

Olkin, R. (2002). Could you hold the door for me? Including disability in diversity. *Cultural Diversity and Ethnic Minority Psychology, 8*(2), 130.

Olusanya, B. O., & Okolo, A. A. (2006). Revisiting the Ten Questions questionnaire for developing countries. *International Journal of Epidemiology, 35*(4), 1103-1103.

Olusanya, B. O., Davis, A. C., Wertlieb, D., Boo, N. Y., Nair, M. K. C., Halpern, R., ... & Kassebaum, N. J. (2018). Developmental disabilities among children younger than 5 years in 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Global Health, 6*(10), e1100-e1121.

O'Neil, M. J., & Jackson, L. (1983). Nominal Group Technique: A process for initiating curriculum development in higher education. *Studies in Higher Education, 8*(2), 129-138.

Ooi, Y. P., Rescorla, L., Ang, R. P., Woo, B., & Fung, D. S. S. (2011). Identification of autism spectrum disorders using the Child Behavior Checklist in Singapore. *Journal of Autism and Developmental Disorders, 41*(9), 1147-1156

Oshodi, Y. O., Olagunju, A. T., Oyelohunnu, M. A., Campbell, E. A., Umeh, C. S., Aina, O. F., Adeyemi, J. D. (2016). Autism spectrum disorder in a community-based sample with neurodevelopmental problems in Lagos, Nigeria. *Journal of Public Health in Africa, 7*(2), 559; 559-559.

Otieno, P. A. (2009). Biblical and theological perspectives on disability: Implications on the rights of persons with disability in Kenya. *Disability Studies Quarterly, 29*(4)

Ouellette-Kuntz, H., Shooshtari, S., Temple, B., Brownell, M., Burchill, C., Yu, C. T., ... & Hennen, B. (2009). Estimating administrative prevalence of intellectual disabilities in Manitoba. *Journal on Developmental Disabilities*, *15*(3), 69.

Ousley, O., & Cermak, T. (2014). Autism spectrum disorder: Defining dimensions and subgroups. *Current Developmental Disorders Reports, 1*(1), 20-28.

Özdemir, O., & Diken, I. H. (2019). Reliability and Validity Studies of the Adapted Autism Behaviour Checklist in Turkey. *Journal of Developmental and Physical Disabilities*, *31*(3), 359-376.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of child Psychology and Psychiatry*, *32*(7), 1081-1105.

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 523-540.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, 105906.

Palmer, M., Larkin, M., de Visser, R., & Fadden, G. (2010). Developing an interpretative phenomenological approach to focus group data. *Qualitative Research in Psychology*, *7*(2), 99-121.

Parmenter, T. R. (2011). What is intellectual disability? How is it assessed and classified?. *International Journal of Disability, Development and Education*, *58*(3), 303-319.

Pasco, G. (2010). Identification and diagnosis of autism spectrum disorders: an update. *Pediatric Health*, *4*(1), 107-114.

Pasquali, L. (2009). Psychometrics. *Revista da Escola de Enfermagem da USP*, *43*, 992-999.

Patel, D. R., Apple, R., Kanungo, S., & Akkal, A. (2018). Narrative review of intellectual disability: definitions, evaluation and principles of treatment. *Pediatr Med*, *1*, 11.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, *78*(4), 1255-1264.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*(1), 51-87.

Pepperdine, C. R., & McCrimmon, A. W. (2018). Test review: Vineland Adaptive Behavior Scales, third edition (Vineland-3) by sparrow, S. S., cicchetti, D. V., & Saulnier, C. A. *Canadian Journal of School Psychology*, *33*(2), 157-163.

Petrocchi, S., Levante, A., & Lecciso, F. (2020). Systematic review of level 1 and level 2 screening tools for autism spectrum disorders in toddlers. *Brain Sciences*, *10*(3), 180.

Posserud, M. B., Lundervold, A. J., & Gillberg, C. (2006). Autistic features in a total population of 7–9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire). *Journal of Child Psychology and Psychiatry*, *47*(2), 167-175.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515-526.

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147–1157.

Prudon, P. (2014). Confirmatory factor analysis: A brief introduction and critique. *Qualtrics, P.UT, USA.*

Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, *27*(2), 224-260.

Rao, V. S., Mysore, A. V., & Raman, V. (2016). The neuropsychology of autism-A focus on three major theories. *Journal of Indian Association for Child and Adolescent Mental Health*, *12*(2), 162-199.

Rapin, I., & Katzman, R. (1998). Neurobiology of autism. *Annals of Neurology*, *43*(1), 7-14.

Rapoport, J., Chavez, A., Greenstein, D., Addington, A., & Gogtay, N. (2009). Autism spectrum disorders and childhood-onset schizophrenia: Clinical and biological contributions to a relation revisited. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(1), 10-18.

Ravindran, N., & Myers, B. J. (2012). Cultural influences on perceptions of health, illness, and disability: A review and focus on autism. *Journal of child and family studies*, *21*(2), 311-319.

Raykov, T., & Marcoulides, G. A. (2011). Classical item analysis using latent variable modeling: A note on a direct evaluation procedure. *Structural Equation Modeling*, *18*(2), 315-324.

Retief, M., & Letšosa, R. (2018). Models of disability: A brief overview. *HTS Teologiese Studies/Theological Studies*, *74*(1)

Roberts, A. R., & Kurtz, L. F. (1987). Historical perspectives on the care and treatment of the mentally ill. *The Journal of Sociology & Social Welfare*, *14*(4), Article 5.

Robichaux, M.A., Cowan, C.W. (2013). Signaling Mechanisms of Axon Guidance and Early Synaptogenesis. In: Andersen, S., Pine, D. (eds) *The Neurobiology of Childhood. Current Topics in Behavioral Neurosciences*, Vol 16. Springer, Berlin, Heidelberg.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *31*(2), 131-144.

Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism*, *12*(5), 537-556.

Rodier, P. M., & Hyman, S. L. (1998). Early environmental factors in autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *4*(2), 121-128.

Rogers, C. L., Goddard, L., Hill, E. L., Henry, L. A., & Crane, L. (2016). Experiences of diagnosing autism spectrum disorder: A survey of professionals in the United Kingdom. *Autism*, *20*(7), 820-831.

Rotatori, A. F., & Epstein, M. (1978). The Slosson Intelligence Test as a quick screening test of mental ability with profoundly and severely retarded children. *Psychological Reports*, 42(3_suppl), 1117-1118.

Rozga, A., Anderson, S., & Robins, D. L. (2011). Major current neuropsychological theories of ASD. *The Neuropsychology of Autism*, , 97-120.

Ruparelia, K., Abubakar, A., Badoe, E., Bakare, M., Visser, K., Chugani, D. C., Chugani, H. T., Donald, K. A., Wilmshurst, J. M., Shih, A., Skuse, D., & Newton, C. R. (2016). Autism spectrum disorders in Africa: Current challenges in identification, assessment, and treatment: A report on the International Child Neurology Association Meeting on ASD in Africa, Ghana, April 3-5, 2014. *Journal of child neurology*, *31*(8), 1018-1026.

Rusch, T., Lowry, P. B., Mair, P., & Treiblmaier, H. (2017). Breaking free from the limitations of classical test theory: Developing and measuring information systems scales using item response theory. *Information & Management*, *54*(2), 189-203.

Rushton, J. P., & Jensen, A. R. (2006). The totality of available evidence shows the race IQ gap still remains. *Psychological Science-Cambridge-*, *17*(10), 921.

Russell, G., Stapley, S., Newlove-Delgado, T., Salmon, A., White, R., Warren, F., Pearson, A., & Ford, T. (2022). Time trends in autism diagnosis over 20 years: A UK population-based cohort study. *Journal of Child Psychology and Psychiatry*, *63*(6), 674-682.

Rutter, M. L. (2011). Progress in understanding autism: 2007–2010. *Journal of Autism and Developmental Disorders*, *41*(4), 395-404.

Rutter, M., Bailey, A., Bolton, P., & Le Couteur, A. (1994). Autism and known medical conditions: Myth and substance. *Journal of Child Psychology and Psychiatry*, *35*(2), 311-322.

Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism Diagnostic Interview-Revised. *Los Angeles, CA: Western Psychological Services*, *29*(2003), 30.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: Preliminary findings. *Journal of Child Psychology and Psychiatry, 8*(1), 1-11.

Saloojee, G., Phohole, M., Saloojee, H., & IJsselmuiden, C. (2007). Unmet health, welfare and educational needs of disabled children in an impoverished South African peri-urban township. *Child Care, Health and Development, 33*(3), 230-235.

Sandin, S., Nygren, K., Iliadou, A., Hultman, C. M., & Reichenberg, A. (2013). Autism and mental retardation among offspring born after in vitro fertilization. *Jama, 310*(1), 75-84.

Sangare, M., Toure, H. B., Toure, A., Karembe, A., Dolo, H., Coulibaly, Y. I., ... & Geschwind, D. H. (2019). Validation of two parent-reported autism spectrum disorders screening tools M-CHAT-R and SCQ in Bamako, Mali. *Eneurologicalsci*, *15*, 100188.

Sango, P. N. (2017). Country profile: Intellectual and developmental disability in Nigeria. *Tizard Learning Disability Rev, 22*(2), 87-93.

Sappok, T., Gaul, I., Bergmann, T., Dziobek, I., Bölte, S., Diefenbacher, A., & Heinrich, M. (2014a). The Diagnostic Behavioral Assessment for Autism Spectrum Disorder—Revised: A screening instrument for adults with intellectual disability suspected of autism spectrum disorders. *Research in Autism Spectrum Disorders, 8*(4), 362-375.

Sappok, T., Gaul, I., Dziobek, I., Bölte, S., Diefenbacher, A., & Bergmann, T. (2014b). Der Diagnostische Beobachtungsbogen für Autismus Spektrumstörungen (DiBAS): Ein Screeninginstrument für Erwachsene mit Intelligenzminderung bei Autismusverdacht. *Psychiatrische Praxis*, 41, 1–7.

Savalli, C., Resende, B., & Gaunet, F. (2016). Eye contact is crucial for referential communication in pet dogs. *PLoS One, 11*(9), e0162161.

Sawaki, Y. (2016). 4. Norm-referenced vs. criterion-referenced approach to assessment. In D. Tsagari & J. Banerjee (Ed.), *Handbook of Second Language Assessment* (pp. 45-60). Berlin, Boston: De Gruyter Mouton.

Sawyer, S. M., Afifi, R. A., Bearinger, L. H., Blakemore, S., Dick, B., Ezeh, A. C., & Patton, G. C. (2012). Adolescence: A foundation for future health. *The Lancet*, 379(9826), 1630-1640.

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. *The Lancet Child & Adolescent Health*, *2*(3), 223-228.

Sawyer, R. N., & Whitten, J. R. (1972). Concurrent validity of the Quick Test. Psychological reports, 30(1), 64-66.

Schalock, R. L. & Luckasson, R. A. (2013) What's at stake in the lives of people with intellectual disability? Part I: The power of naming, defining, diagnosing, classifying, and planning supports. *Intellectual and Developmental Disabilities*, 51(2), 86-93.

Schalock, R. L., Luckasson, R. A., & Shogren, K. A. (2007). The renaming of mental retardation: Understanding the change to the term intellectual disability. *Intellectual and Developmental Disabilities, 45*(2), 116-124.

Schanding, G. T., Nowell, K. P., & Goin-Kochel, R. P. (2012). Utility of the Social Communication Questionnaire-current and Social Responsiveness Scale as teacher-report screening tools for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(8), 1705-1716.

Scherzer, A. L., Chhagan, M., Kauchali, S., & Susser, E. (2012). Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. *Developmental Medicine & Child Neurology, 54*(12), 1079-1084.

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.

Sergeant, J. (2000). The cognitive-energetic model: An empirical approach to attention-deficit hyperactivity disorder. *Neuroscience & Biobehavioral Reviews*, *24*(1), 7-12.

Shakes, P., & Cashin, A. (2019). Identifying language for people on the autism spectrum: A scoping review. *Issues in Mental Health Nursing*, *40*(4), 317-325.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist*, *44*(6), 922.

Shell-Duncan, B. (2008). From health to human rights: Female genital cutting and the politics of intervention. *American anthropologist*, *110*(2), 225-236.

Shell-Duncan, B., Hernlund, Y., Wander, K., & Moreau, A. (2013). Legislating Change? Responses to Criminalizing Female Genital Cutting in Senegal. *Law & society review*, *47*(4), 803–835.

Shepherd, C. A., & Waddell, C. (2015). A qualitative study of autism policy in Canada: Seeking consensus on children's services. *Journal of Autism and Developmental Disorders*, *45*(11), 3550-3564.

Shree, A., & Shukla, P. C. (2016). Intellectual Disability: definition, classification, causes and characteristics. *Learning Community-An International Journal of Educational and Social Development*, *7*(1), 9-20.

Shuttleworth-Edwards, B. A., Donnelly, J. R., Martin, I. R., & Radloff, E. S. (2004). A cross-cultural study with culture fair normative indications on WAIS-III Digit Symbol—Incidental learning. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 921-932.

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G., Hartman, N. P., & Radloff, S. E. (2004). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAIS-III test performance. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 903-920.

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., ... & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences*, *111*(42), 15220-15225.

Skirrow, C., Cashdollar, N., Granger, K., Jennings, S., Baker, E., Barnett, J., & Cormack, F. (2022). Test-retest reliability on the Cambridge neuropsychological test automated battery: Comment on karlsen et al.(2020). *Applied Neuropsychology: Adult*, *29*(5), 889-892.

Slaughter, V. (2015). Theory of mind in infants and young children: A review. *Australian Psychologist*, *50*(3), 169-172.

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.

Smith, L. (2015). *Afrikaans Autism Diagnostic Observation Schedule-2: translation and cultural appropriateness for the coloured population from low-middle socioeconomic backgrounds living in the Western Cape* (Master's thesis, University of Cape Town).

Smith, N. (2011). The face of disability in Nigeria: A disability survey in Kogi and Niger States. *Disability, CBR & Inclusive Development*, *22*(1), 35-47.

Snow, A. V., & Lecavalier, L. (2008). Sensitivity and specificity of the Modified Checklist for Autism in Toddlers and the Social Communication Questionnaire in preschoolers suspected of having pervasive developmental disorders. *Autism*, *12*(6), 627-644.

Snow, A. (2013). Social communication questionnaire. *Encyclopedia of Autism Spectrum Disorders* (pp. 2893-2895). Springer.

Soleimani F., Khakshour A., Abasi Z., Khayat S., Ghaemi S. Z., Golchin N. A. H. Review of autism screening tests. *International Journal of Pediatrics.* 2014;2(4):319–329.

Solmi, M., Song, M., Yon, D. K., Lee, S. W., Fombonne, E., Kim, M. S., … & Cortese, S. (2022). Incidence, prevalence, and global burden of autism spectrum disorder from 1990 to 2019 across 204 countries. *Molecular Psychiatry*, 1-9.

Søndenaa, E., Bjørgen, T. G., & Nøttestad, J. A. (2007). Validation of the Norwegian version of Hayes Ability Screening Index for mental retardation. *Psychological reports, 101(3),* 1023–1030.

Søndenaa, E., Nygård, Ø., Nøttestad, J. A., & Linaker, O. M. (2011). Validation and adaptation of the Norwegian version of Hayes Ability Screening Index for intellectual difficulties in a psychiatric sample. *Nordic Journal of Psychiatry, 65*(1), 47-51.

Søndenaa, E., Rasmussen, K., Palmstierna, T., & Nøttestad, J. (2008). The prevalence and nature of intellectual disability in Norwegian prisons. *Journal of Intellectual Disability Research, 52*(12), 1129-1137.

Søndergaard, E., Ertmann, R. K., Reventlow, S., & Lykke, K. (2018). Using a modified Nominal Group Technique to develop general practice. *BMC Family Practice, 19*(1), 1-9.

Soto, S., Linas, K., Jacobstein, D., Biel, M., Migdal, T., & Anthony, B. J. (2015). A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism*, 19(6), 646-661.

Sparrow, S., Cicchetti, D., & Saulnier, C. (2016). *Vineland Adaptive Behavior Scales–Third Edition (Vineland-3)*. Circle Pines, MN: American Guidance Service.

Steel, M. J. G., Gorman, R., & Flexman, J. E. (1984). Neuropsychiatric testing in an autistic mathematical idiot-savant: Evidence for nonverbal abstract capacity. *Journal of the American Academy of Child Psychiatry, 23*(6), 704-707.

Stein, Z., Durkin, M., & Belmont, L. (1986). "Serious" mental retardation in developing countries: An epidemiologic approach. *Annals of the New York Academy of Sciences, 477*(1), 8-21.

Steiner, A. M., Goldsmith, T. R., Snow, A. V., & Chawarska, K. (2012). Practitioner's guide to assessment of autism spectrum disorders in infants and toddlers. *Journal of Autism and Developmental Disorders, 42*(6), 1183-1196.

Stephens, M. (2012). Screening for autism spectrum disorders in South Africa: Using the Modified Checklist for Autism in Toddlers (M-CHAT). *(Unpublished Honours Thesis). University of Cape Town, Cape Town, South Africa.*

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*(2684), 677-680.

Stirk, S., Field, B., & Black, J. (2018). An independent investigation of the utility of the Learning Disability Screening Questionnaire (LDSQ) within a community learning disability team. *Journal of Applied Research in Intellectual Disabilities, 31*(2), e223-e228.

Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Brief report: screening tool for autism in two-year-olds (STAT): development and preliminary data. *Journal of autism and developmental disorders, 30*(6), 607.

Streiner, D. L., & Cairney, J. (2007). What's under the ROC? An introduction to receiver operating characteristics curves. *The Canadian Journal of Psychiatry, 52*(2), 121-128.

Suhail, K., & Zafar, F. (2008). Prevalence of autism in special education schools of Lahore. *Pakistan Journal of Psychological Research, 23*(2).

Swinkels, S. H., Dietz, C., van Daalen, E., Kerkhof, I. H., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum in children aged 14 to 15 months. I: The development of the Early Screening of Autistic Traits questionnaire (ESAT). *Journal of Autism and Developmental Disorders, 36*(6), 723-732.

Tammela, O. (2013). Applications of consensus methods in the improvement of care of paediatric patients: A step forward from a 'good guess'. *Acta Paediatrica, 102*(2), 111-115.

Tassé, M. J., Schalock, R. L., Balboni, G., Bersani Jr, H., Borthwick-Duffy, S. A., Spreat, S., ... & Zhang, D. (2012). The construct of adaptive behavior: Its conceptualization, measurement, and use in the field of intellectual disability. *American journal on intellectual and developmental disabilities*, *117*(4), 291-303.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34-42.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018a). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research,* 27(5), 1159–1170.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., De Vet, H., Bouter, L. M., Alonso, J., Westerman. M. J., Patrick, D. L., & Mokkink, L. B. (2018b). COSMIN methodology for assessing the content validity of PROMs–user manual. *Amsterdam: VU University Medical Center,* Netherlands.

Thabtah, F., & Peebles, D. (2019). Early autism screening: A comprehensive review. *International Journal of Environmental Research and Public Health*, *16*(18), 3502.

To, W. T., Vanheule, S., Vanderplasschen, W., Audenaert, K., & Vandevelde, S. (2015). Screening for intellectual disability in persons with a substance abuse problem: Exploring the validity of the Hayes Ability Screening Index in a Dutch-speaking sample. *Research in Developmental Disabilities, 36,* 498-504.

Tomkins, L., & Eatough, V. (2010). Reflecting on the use of IPA with focus groups: Pitfalls and potentials. *Qualitative Research in Psychology*, *7*(3), 244-262.

Tomlinson, M., Yasamy, M. T., Emerson, E., Officer, A., Richler, D., & Saxena, S. (2014). Setting global research priorities for developmental disabilities, including intellectual disabilities and autism. *Journal of Intellectual Disability Research, 58*(12), 1121-1130.

Towle, P. O., & Patrick, P. A. (2016). Autism spectrum disorder screening instruments for very young children: A systematic review. *Autism Research and Treatment, 2016*, 4624829.

Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health, 5,* 307.

Trivedi, A. (1977). A comparison of three intelligence tests for the assessment of mental retardation. *Journal of Mental Deficiency Research*, *21*(4), 289-297.

Uljarević, M., Frazier, T. W., Phillips, J. M., Jo, B., Littlefield, S., & Hardan, A. Y. (2021). Quantifying research domain criteria social communication subconstructs using the Social Communication Questionnaire in youth. *Journal of Clinical Child & Adolescent Psychology*, *50*(5), 609-618.

Ung, D., Johnco, C., McBride, N. M., Howie, F., Scalli, L., & Storch, E. A. (2016). Optimizing the screening of autism spectrum disorders in outpatient clinics: An examination of the Social Communication Questionnaire-Lifetime. *Research in Autism Spectrum Disorders, 27,* 21-28.

UNICEF. (2014). Generation 2030 Africa: Child demographics in Africa. Division of Data. *Research Policy.* Retrieved from https://data.unicef.org/resources/generation-2030-africa-child-demographics-in-africa/ Accessed 4 February 2022.

United Nations Department of Economic and Social Affairs. (2021). *World Economic Situation and Prospects 2021*. United Nations Publications.

United Nations, Department of Economic and Social Affairs, Population Division. (2017). *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. Working Paper no. ESA/P/WP/248. (accessed January 25, 2019).*https://population.un.org/wpp/Publications/Files/WPP2017_KeyFindings.pdf

United Nations. (2014). Country classification: Data sources, country classifications and aggregation methodology. *World Economic Situation and Prospects 2014*, 143-150.

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., De-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: predictive coding in autism. *Psychological Review*, *121*(4), 649.

Van de Ven, A. H., & Delbecq, A. L. (1972). The Nominal Group as a research instrument for exploratory health studies. *American Journal of Public Health*, *62*(3), 337-342.

Van de Vijver, Fons. J. R., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, *13*(1), 29-37.

Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology, 54*(2), 119-135.

Van der Linde, J., Swanepoel, D. W., Glascoe, F. P., Louw, E. M., & Vinck, B. (2015). Developmental screening in South Africa: Comparing the national developmental checklist to a standardized tool. *African Health Sciences*, 15(1), 188-196.

Van Krevelen, D. A. (1971). Early infantile autism and autistic psychopathy. *Journal of Autism and Childhood Schizophrenia, 1*(1), 82-86.

Van Ryn, M., & Heaney, C. A. (1992). What's the use of theory? *Health Education Quarterly, 19*(3), 315-330.

Vawda, N., Milburn, N. G., Steyn, R., & Zhang, M. (2017). The development of a screening tool for the early identification of risk for suicidal behavior among students in a developing country. *Psychological Trauma: Theory, Research, Practice, and Policy, 9*(3), 267.

Vivanti, G., & Messinger, D. S. (2021). Theories of autism and autism treatment from the DSM III through the present and beyond: Impact on research and practice. *Journal of Autism and Developmental Disorders, 51*(12), 4309-4320.

Vock. (2008). Gaussian Distribution of Intelligence | IHVO. Retrieved March 1, 2023, from http://www.ihvo.de/202/gaussian-distribution-of-intelligence/

Volker, M. A., & Lopata, C. (2008). Autism: A review of biological bases, assessment, and intervention. *School Psychology Quarterly*, *23*(2), 258.

Volkmar, F., Siegel, M., Woodbury-Smith, M., King, B., McCracken, J., & State, M. (2014). Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(2), 237-257.

Vrancic, D., Nanclares, V., Soares, D., Kulesz, A., Mordzinski, C., Plebst, C., & Starkstein, S. (2002). Sensitivity and specificity of the Autism Diagnostic Inventory-Telephone Screening in Spanish. *Journal of Autism & Developmental Disorders, 32*(4), 313-320

Wall, D. P., Dally, R., Luyster, R., Jung, J. Y., & Deluca, T. F. (2012). Use of artificial intelligence to shorten the behavioral diagnosis of autism. *PLoS ONE*, 7(8).

Wallsten, T. S., & Budescu, D. V. (1983). State of the art—Encoding subjective probabilities: A psychological and psychometric review. *Management Science*, *29*(2), 151-173.

Watson, L. R., Baranek, G. T., Crais, E. R., Reznick, J. S., Dykstra, J., & Perryman, T. (2007). The first year inventory: Retrospective parent responses to a questionnaire designed to identify one-year-olds at risk for autism. *Journal of Autism and Developmental Disorders, 37(1),* 49-61.

Webb, E., Morey, J., Thompsen, W., Butler, C., Barber, M., & Fraser, W. (2003). Prevalence of autistic spectrum disorder in children attending mainstream schools in a Welsh education authority. *Developmental Medicine & Child Neurology, 45*(6), 377-384.

Webb, N. M., & Shavelson, R. J. (2005). Generalizability theory: Overview. In Brian S. Everitt & David C. Howell (Ed.), *Encyclopedia of Statistics in Behavioral Science* (Volume 2, pp. 717–719). John Wiley & Sons, Ltd, Chichester.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children (5th ed.)*. Bloomington, MN: NCS Pearson.

Wei, T., Chesnut, S. R., Barnard-Brak, L., & Richman, D. (2015). Psychometric analysis of the Social Communication Questionnaire using an item-response theory framework: Implications for the use of the lifetime and current forms. *Journal of Psychopathology and Behavioral Assessment, 37*(3), 469-480.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*, *19*(1), 231-240.

Westerlund, M., & Sundelin, C. (2000). Screening for developmental language disability in 3-year-old children. experiences from a field study in a Swedish municipality. *Child Care, Health and Development Original Articles*, *26*(2), 91-110.

Weye, N., Santomauro, D. F., Agerbo, E., Christensen, M. K., Iburg, K. M., Momen, N. C., ... & Plana-Ripoll, O. (2021). Register-based metrics of years lived with disability associated with mental and substance use disorders: a register-based cohort study in Denmark. *The Lancet Psychiatry*, *8*(4), 310-319.

White, H., & Waddington, H. (2012). Why do we care about evidence synthesis? an introduction to the special issue on systematic reviews. *Journal of Development Effectiveness*, *4*(3), 351-358.

Wicherts, J. M., Dolan, C. V., & van der Maas, Han LJ. (2010). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence*, *38*(1), 1-20.

Wild, D., Furtado, T., & Angalakuditi, M. (2012). The translation and cultural adaptation of the Child Behavior Checklist for use in Israel (Hebrew), Korea, the US (Spanish), India (Malayalam and Kannada), and Spain. *Psychology Research and Behavior Management*, *5*, 51-56.

Williams, D. L., Goldstein, G., & Minshew, N. J. (2006). Neuropsychologic functioning in children with autism: Further evidence for disordered complex information-processing. *Child Neuropsychology*, *12*(4-5), 279-298.

Williams, J. G., Higgins, J. P., & Brayne, C. E. (2006). Systematic review of prevalence studies of autism spectrum disorders. *Archives of Disease in Childhood*, *91*(1), 8-15.

Williamson, P. R., Altman, D. G., Blazeby, J. M., Clarke, M., Devane, D., Gargon, E., & Tugwell, P. (2012). Developing core outcome sets for clinical trials: Issues to consider. *Trials*, *13*(1), 1-8.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*(1), 11-29.

Wong, H. B., & Lim, G. H. (2011). Measures of diagnostic accuracy: Sensitivity, specificity, PPV and NPV. *Proceedings of Singapore Healthcare, 20*(4), 316-318.

World Health Organisation Africa. (2023). Barriers to mental health care in Africa. Retrieved February 1, 2023, from https://www.afro.who.int/news/barriers-mental-health-care-africa

World Bank. (n.d.). Population, total - Nigeria | Data. Retrieved February 7, 2023, from https://data.worldbank.org/indicator/SP.POP.TOTL?locations=NG

World Bank. (2020). World Bank country and lending groups. Retrieved from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519. Accessed 31 Mar 2022.

World Health Organization. (2001). *Towards a common language for functioning, disability and health. international classification of functioning, disability and health (ICF).*http://www.who.int/classifications/icf/en/

World Health Organization. (2014). *Health for the World's Adolescents: A Second Chance in the Second Decade: Summary.* Retrieved from https://apps.who.int/iris/handle/10665/112750. Accessed 4 February 2022.

World Health Organization. (2015). *WHO global disability action plan 2014-2021: Better health for all people with disability*, World Health Organization.

World Health Organization. (2017). *International classification of functioning, disability and health (ICF).*https://apps.who.int/classifications/icfbrowser/

World Health Organization. (2020). *International Statistical Classification of Diseases and Related Health Problems (ICD) (11th Edition)*. https://icd.who.int/browse11/l-m/en. Accessed 12 November 2020.

Yasuda, H., Yoshida, K., Yasuda, Y., & Tsutsui, T. (2011). Infantile zinc deficiency: association with autism spectrum disorders. *Scientific reports*, *1*(1), 129.

Yates, K., & Le Couteur, A. (2016). Diagnosing autism/autism spectrum disorders. *Paediatrics and Child Health*, *26*(12), 513-518.

Yesufu, M. L. (2016). The impact of religion on a secular state: The Nigerian experience. *Studia Historiae Ecclesiasticae, 42*(1), 1-11.

Young, R. (2007). *Autism detection in early childhood (ADEC) manual*. Camberwell: ACER Press.

Zheng, S., LeWinn, K., Ceja, T., Hanna-Attisha, M., O'Connell, L., & Bishop, S. (2021). Adaptive behavior as an alternative outcome to intelligence quotient in studies of children at risk: A study of preschool-aged children in flint, MI, USA. *Frontiers in Psychology, 12*, 692330.

Zheng, W., Eilam-Stock, T., Wu, T., Spagna, A., Chen, C., Hu, B., & Fan, J. (2019). Multi-feature based network revealing the structural abnormalities in autism spectrum disorder. *IEEE Transactions on Affective Computing*, *12*(3), 732-742.

Appendices

Appendix 1 – Tizard Ethics Approval

TIZARD
University of Kent

Tizard Ethics Feedback Form

Student Name:	Eziafakaku Nwokolo (Ziffy)
Supervisor:	Glynis Murphy & Peter Langdon
Title:	*The reliability and validity of screening tools for Intellectual and Developmental Disabilities and Autism in adolescents in Nigeria.*
The Chair of the Tizard Ethics Committee has considered the recent amendments submitted to the above proposal and confirms that this has ethical approval. **Signed:** **J.Ruffels** **Date: 04.07.19** **On behalf of Tizard Ethics Committee**	
Alterations approved by Supervisor	**Signature** **Date** 30/05/19
Final approval On behalf of Tizard Ethics Committee	**Michelle McCarthy** **Signature** **Date 3.7.19**

Appendix 2 – Collaborative Institutional Training Initiative (CITI) completion certificate

COLLABORATIVE INSTITUTIONAL TRAINING INITIATIVE (CITI PROGRAM)
COMPLETION REPORT - PART 1 OF 2
COURSEWORK REQUIREMENTS*

* NOTE: Scores on this Requirements Report reflect quiz completions at the time all requirements for the course were met. See list below for details. See separate Transcript Report for more recent quiz scores, including those on optional (supplemental) course elements.

- Name: Ezialakaku Nwokolo (ID: 7902492)
- Institution Affiliation: Independent Learner (ID: 569)
- Phone: 8035351637

- Curriculum Group: Human Subjects Research - BASIC
- Course Learner Group: Human Subjects Research – Social-Behavioral-Educational Basic
- Stage: Stage 1 - Independent Learner

- Record ID: 30562232
- Completion Date: 16-Feb-2019
- Expiration Date: 16-Feb-2020
- Minimum Passing: 80
- Reported Score*: 91

REQUIRED AND ELECTIVE MODULES ONLY	DATE COMPLETED	SCORE
Conflicts of Interest in Human Subjects Research (ID: 17464)	13-Feb-2019	5/5 (100%)
History and Ethical Principles - SBE (ID: 490)	15-Feb-2019	4/5 (80%)
Defining Research with Human Subjects - SBE (ID: 491)	15-Feb-2019	4/5 (80%)
The Federal Regulations - SBE (ID: 502)	16-Feb-2019	5/5 (100%)
Assessing Risk - SBE (ID: 503)	16-Feb-2019	4/5 (80%)
Informed Consent - SBE (ID: 504)	16-Feb-2019	5/5 (100%)
Privacy and Confidentiality - SBE (ID: 505)	16-Feb-2019	4/5 (80%)
Unanticipated Problems and Reporting Requirements in Social and Behavioral Research (ID: 14928)	16-Feb-2019	5/5 (100%)
Populations in Research Requiring Additional Considerations and/or Protections (ID: 16680)	16-Feb-2019	5/5 (100%)

For this Report to be valid, the learner identified above must have had a valid affiliation with the CITI Program subscribing institution identified above or have been a paid Independent Learner.

Verify at: www.citiprogram.org/verify/?x5f9c701d-ad2c-438a-9f61-53a1bcfebea5-30562232

Collaborative Institutional Training Initiative (CITI Program)
Email: support@citiprogram.org
Phone: 888-529-5929
Web: https://www.citiprogram.org

COLLABORATIVE INSTITUTIONAL TRAINING INITIATIVE (CITI PROGRAM)
COMPLETION REPORT - PART 2 OF 2
COURSEWORK TRANSCRIPT**

** NOTE: Scores on this Transcript Report reflect the most current quiz completions, including quizzes on optional (supplemental) elements of the course. See list below for details. See separate Requirements Report for the reported scores at the time all requirements for the course were met.

- Name: Ezialakaku Nwokolo (ID: 7902492)
- Institution Affiliation: Independent Learner (ID: 569)
- Phone: 8035351637

- Curriculum Group: Human Subjects Research - BASIC
- Course Learner Group: Human Subjects Research – Social-Behavioral-Educational Basic
- Stage: Stage 1 - Independent Learner

- Record ID: 30562232
- Report Date: 16-Feb-2019
- Current Score**: 91

REQUIRED, ELECTIVE, AND SUPPLEMENTAL MODULES	MOST RECENT	SCORE
Populations in Research Requiring Additional Considerations and/or Protections (ID: 16680)	16-Feb-2019	5/5 (100%)
Defining Research with Human Subjects - SBE (ID: 491)	15-Feb-2019	4/5 (80%)
The Federal Regulations - SBE (ID: 502)	16-Feb-2019	5/5 (100%)
Assessing Risk - SBE (ID: 503)	16-Feb-2019	4/5 (80%)
Informed Consent - SBE (ID: 504)	16-Feb-2019	5/5 (100%)
Privacy and Confidentiality - SBE (ID: 505)	16-Feb-2019	4/5 (80%)
Unanticipated Problems and Reporting Requirements in Social and Behavioral Research (ID: 14928)	16-Feb-2019	5/5 (100%)
History and Ethical Principles - SBE (ID: 490)	15-Feb-2019	4/5 (80%)
Conflicts of Interest in Human Subjects Research (ID: 17464)	13-Feb-2019	5/5 (100%)

For this Report to be valid, the learner identified above must have had a valid affiliation with the CITI Program subscribing institution identified above or have been a paid Independent Learner.

Verify at: www.citiprogram.org/verify/?x5f9c701d-ad2c-438a-9f61-53a1bcfebea5-30562232

Collaborative Institutional Training Initiative (CITI Program)
Email: support@citiprogram.org
Phone: 888-529-5929
Web: https://www.citiprogram.org

Appendix 3 – National Health Research Ethics Committee of Nigeria (NHREC)

National Health Research Ethics Committee of Nigeria (NHREC)

Promoting Highest Ethical and Scientific Standards for Health Research in Nigeria

Federal Ministry of Health

NHREC Protocol Number NHREC/01/01/2007-30/07/2019
HREC Approval Number NHREC/01/01/2007-16/09/2019

Date: 16 September, 2019

Re: The Reliability & Validity of Screening Tools for Intellectual Disabilities and autism Spectrum Disorder in Nigerian Adolescent

Health Research Committee assigned number: NHREC/01/01/2007
Name of Student Investigator: Eziafakaku Nwokolo
Address of Student Investigator: Tizard Center
Faculty of Social Science
University of Kent
Canterbury, United Kingdom
Email: eziafakaku@gmail.com
Tel: +2348035351637
Date of receipt of valid application: 30/07/2019
Date when final determination of research was made: 14-09-2019

Notice of Expedited Committee Review and Approval

This is to inform you that the research described in the submitted protocol, the consent forms, advertisements and other participant information materials have been reviewed and *given expedited committee approval by the National Health Research Ethics Committee.*

This approval dates from 16/09/2019 to 15/09/2020. If there is delay in starting the research, please inform the HREC so that the dates of approval can be adjusted accordingly. Note that no participant accrual or activity related to this research may be conducted outside of these dates. *All informed consent forms used in this study must carry the HREC assigned number and duration of HREC approval of the study.* In multiyear research, endeavour to submit your annual report to the HREC early in order to obtain renewal of your approval and avoid disruption of your research.

The National Code for Health Research Ethics requires you to comply with all institutional guidelines, rules and regulations and with the tenets of the Code including ensuring that all

adverse events are reported promptly to the HREC. No changes are permitted in the research without prior approval by the HREC except in circumstances outlined in the Code. The HREC reserves the right to conduct compliance visit to your research site without previous notification.

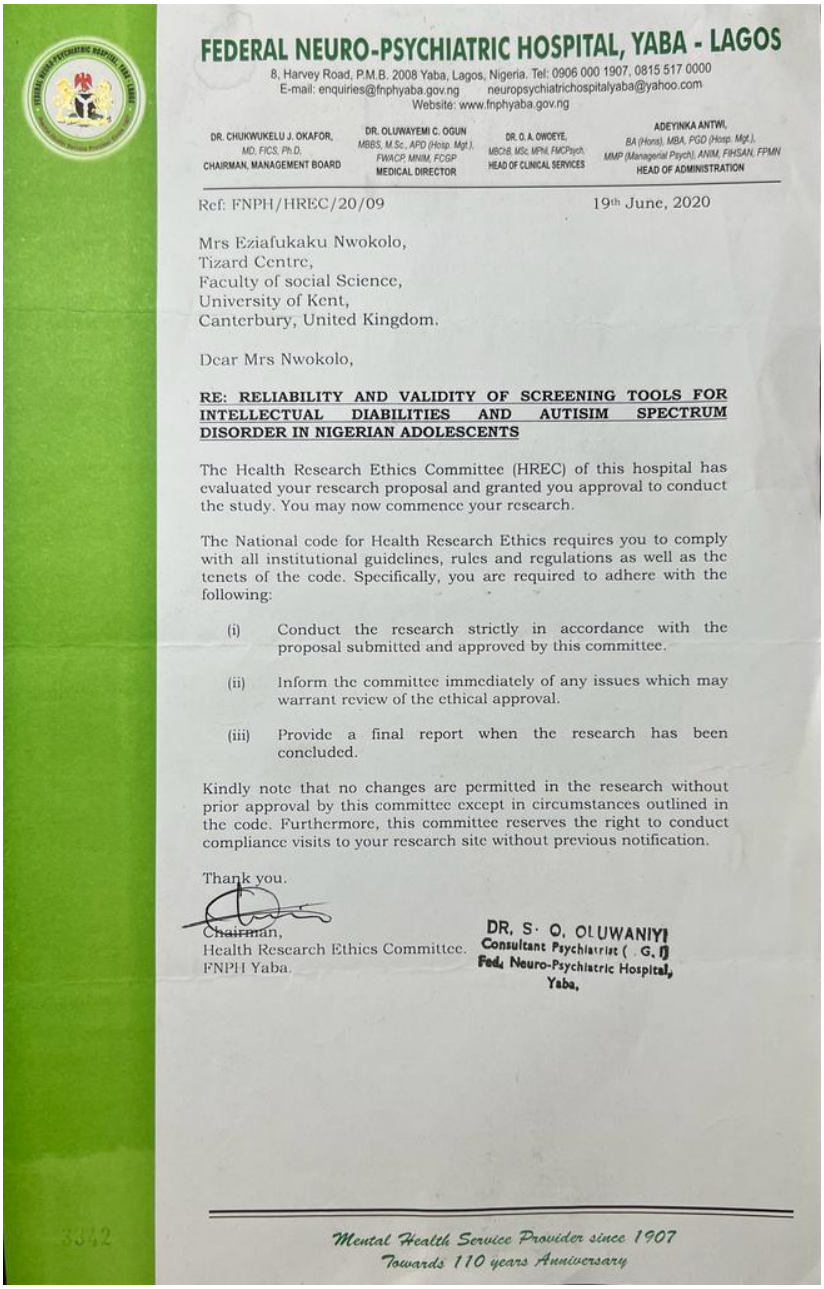
Signed

Professor Zubairu Iliyasu MBBS (UniMaid), MPH (Glasg.), PhD (Shef.), FWACP, FMCPH
Chairman, National Health Research Ethics Committee of Nigeria (NHREC)

Department of Health Planning, Research & Statistics
Federal Ministry of Health
11th Floor, Federal Secretariat Complex Phase III
Ahmadu Bello Way, Abuja

Tel: +234-09-523-8367
E-mail: chairman@nhrec.net, secretary@nhrec.net,
deskofficer@nhrec.net,
URL: http://www.nhrec.net,

Appendix 4 – Federal Neuropsychiatry Hospital Yaba, Lagos Nigeria



FEDERAL NEURO-PSYCHIATRIC HOSPITAL, YABA - LAGOS

8, Harvey Road, P.M.B. 2008 Yaba, Lagos, Nigeria. Tel: 0906 000 1907, 0815 517 0000
E-mail: enquiries@fnphyaba.gov.ng neuropsychiatrichospitalyaba@yahoo.com
Website: www.fnphyaba.gov.ng

DR. CHUKWUKELU J. OKAFOR, MD, FICS, Ph.D, CHAIRMAN, MANAGEMENT BOARD
DR. OLUWAYEMI C. OGUN MBBS, M.Sc., APD (Hosp. Mgt.), FWACP, MNIM, FCGP MEDICAL DIRECTOR
DR. O. A. OWOEYE, MBChB, MSc, MPhil, FMCPsych. HEAD OF CLINICAL SERVICES
ADEYINKA ANTWI, BA (Hons), MBA, PGD (Hosp. Mgt.), MMP (Managerial Psych), ANIM, FIHSAN, FPMN HEAD OF ADMINISTRATION

Ref: FNPH/HREC/20/09 19th June, 2020

Mrs Eziafukaku Nwokolo,
Tizard Centre,
Faculty of social Science,
University of Kent,
Canterbury, United Kingdom.

Dear Mrs Nwokolo,

RE: RELIABILITY AND VALIDITY OF SCREENING TOOLS FOR INTELLECTUAL DIABILITIES AND AUTISIM SPECTRUM DISORDER IN NIGERIAN ADOLESCENTS

The Health Research Ethics Committee (HREC) of this hospital has evaluated your research proposal and granted you approval to conduct the study. You may now commence your research.

The National code for Health Research Ethics requires you to comply with all institutional guidelines, rules and regulations as well as the tenets of the code. Specifically, you are required to adhere with the following:

(i) Conduct the research strictly in accordance with the proposal submitted and approved by this committee.

(ii) Inform the committee immediately of any issues which may warrant review of the ethical approval.

(iii) Provide a final report when the research has been concluded.

Kindly note that no changes are permitted in the research without prior approval by this committee except in circumstances outlined in the code. Furthermore, this committee reserves the right to conduct compliance visits to your research site without previous notification.

Thank you.

Chairman,
Health Research Ethics Committee.
FNPH Yaba.

DR, S. O, OLUWANIYI
Consultant Psychiatrist (G. I)
Fed, Neuro-Psychiatric Hospital,
Yaba,

3342

Mental Health Service Provider since 1907
Towards 110 years Anniversary

Appendix 5 – Information sheet for the Focus Group

Professional group for deciding on best screening tool

Information Sheet for Consensus Meeting

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

To Identify and agree on which screening tools to be used for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents

I am Eziafakaku Nwokolo, a postgraduate student in Intellectual Disabilities at the University of Kent, UK. As part of fulfilling my course requirements, I am conducting a research study on the above topic and would appreciate your voluntary participation in a consensus nominal group meeting. For you to do so, some background information has been provided below.

Purpose of the study

To examine and agree on the contents of which screening tools identified through a systematic review can be used for (a) Intellectual Disability and (b) Autism Spectrum Disorder among Nigerian adolescents.

Why am I being asked to participate?

You have been identified as a parent/carer or a member of one of the relevant professions: paediatrician, psychologist, psychiatrist, teacher or clinician

Must I participate?

It is completely voluntary. If you decide to volunteer your time, then you will keep this information sheet and be given a consent form to sign. You are free to withdraw your participation at any time without giving a reason.

What happens if I decide to participate?

You will be given different measures to review and make comments and recommendations as required. The group meeting should last no longer than a couple of hours and at most, an afternoon. You will be one of the 4 – 6 member group of professionals drawn from paediatricians, child psychiatrists, paediatric neurologists and clinical psychologists. The group's goal will be to provide qualitative information following a series of reviews and discussion about the screening

measures you will be provided. Each of you will independently assess the tools and then suggestions collated. A couple of rounds may be required until there is a consensus on which measures will be more appropriate for use in the Nigerian context. The meeting will be facilitated by myself.

What about my privacy?

Only my supervisors and I will see the meeting notes, and all personal information will be anonymised. Notes will be scanned and saved in an encrypted folder on the computer, while audio recordings will be downloaded to a secure folder on a computer and encrypted.

What happens if I do not like something that's happened?

You will be given a form at the end of your involvement, so that you can give feedback or make a complaint. In addition, at any time, you can contact my supervisors, Prof Peter Langdon by email P.E.Langdon@kent.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk. Alternatively, you can contact the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373. E.Lukehurst@kent.ac.uk

Thank you for your time.

Yours sincerely,

Eziafakaku Nwokolo

Appendix 6 – Consent form for the Focus Group

TIZARD
University of Kent

Professional group for deciding on best screening tool

CONSENT FORM for Consensus Group Meeting

Title: **To Identify and agree on which screening tools to be used for Intellectual Disabilities and Autism in Nigerian adolescents.**

Researcher: Eziafakaku Nwokolo
Email and telephone contact: eun5@kent.ac.uk; +447526316452

Supervisors: Prof Peter Langdon - P.E.Langdon@kent.ac.uk and
Prof Glynis Murphy G.H.Murphy@kent.ac.uk

	Please tick to confirm
• I confirm that I have read and understood the information sheet attached for the above study.	☐
• I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.	☐
• My questions have been answered to my satisfaction.	☐
• I understand that information I give may be shared within the research team and that it will be kept securely.	☐
• I understand that my name will not be used in any publications or presentations arising from the research.	☐
• I agree to take part in the above research study.	☐
• I have read and understood the consent form and agree to	☐

participate.

- I have read and understood the consent form and agree to participate. ☐
- I understand the meeting will be recorded. ☐

Name of Participant:_____ Signature:_____

Date:_____

Appendix 7 – Information sheet for parents, carers and guardians of those with intellectual disability below 18-years

For parents, carers & guardians of children already known to have ID

Information Sheet

Information for parents, carers, and guardians of participants below 18-yrs.

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents

I am Eziafakaku Nwokolo, a postgraduate student in Intellectual Disabilities at the University of Kent, UK. As part of fulfilling my course requirements, I am conducting a research on the above topic and would appreciate your voluntary participation. For you to do so, some background information has been provided below.

Purpose of the study

This research project is aimed at validating a screening tool for intellectual disability in adolescents. Currently, there is no such tool in Nigeria. The process is in two parts and will (a) involve your child responding to some questions on a questionnaire. This should take between 10 – 20 minutes to complete and (b) your child being administered with an intelligence quotient test which will take about 1 hour. We hope that if the research shows that the screening tools are useful, they will become available for doctors, teachers and other health care professionals to use.

What is intellectual disability?

Intellectual disability simply put means, a serious reduction in a person's ability to understand new or complicated information and to learn and apply new skills (impaired intelligence). The result is a decreased ability to look after oneself independently (impaired social functioning) and begins before adulthood.

Why have you been given this sheet?

I would like your child/ward to participate in this research but recognise that s/he is not old enough to give their consent. As such, I am asking you either as a parent, carer or guardian the minor, who is either

(a) not meeting set targets,
(b) performing at a level below his/her peers,
(c) appears to be slow,
(d) is currently residing in/attending a special care home/centre or
(e) has been recommended as a participant by your primary health care provider

to act on their behalf.

Must s/he participate?

It is completely voluntary. If you decide to allow your child/ward to participate, then you will keep this information sheet and be given a consent form to sign. You are free to withdraw your child/ward's participation at any time without giving a reason.

What happens if I allow my child/ward to participate?

Your child/ward will be given a questionnaire to complete. Completion should not take longer than 10 minutes. I will be available to help if s/he needs it. Upon completion, your child/ward will be given an intelligence quotient test to take. This should take about 1 hour. The test will be done by another member of the team that is trained to do so.

Will my child/ward receive treatment or intervention as a result of participating?

This is a research project and no treatment will be given. However, if as a result of the screening your child/ward is identified as having intellectual disability, you will be advised to see your child/ward's health care provider or the referring practitioner.

What about my child/ward's privacy?

Only my supervisors and I will see the returned questionnaires, and all personal information will be anonymised. Filled forms will be scanned and saved in an encrypted folder on the computer and hard copies of the questionnaires will be kept in a locked safe.

What happens if either your child/ward or yourself does not like something that's happened?

You can discuss with myself however, you will be given a form at the end of your child/ward's participation, for feedback or making a complaint. In addition, at any time, you can contact my supervisors, Prof Peter Langdon by email Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk. Alternatively, you can contact the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373. E.Lukehurst@kent.ac.uk

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Appendix 8 – Information sheet for participants with intellectual disability above 18 years

Participants 18yrs and above who have ID

Information Sheet

Information for participants 18-years and above

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents

I am Eziafakaku Nwokolo, a PhD research student.

Purpose of the study

- I want to tell you about some research I am doing.
- Research is a way to learn more about something.
- I would like to find out more about intellectual disability.
- Some people with intellectual disability find it hard to do their work.

Why am I being asked to participate?

- I'm asking you to take part in my research because you have intellectual disability.
- I have spoken to your doctor or the person that takes care or you at home or in the boarding house you live in.

Must I participate?

- You do not have to take part. It is up to you.
- You can say yes now, and you can also say no.
- If you say yes, you can change your mind later.
- If you want to stop, all you have to do is tell me you want to stop.
- I will not be upset at you if you don't want to be in the research.


- If you want to take part, you will keep this information sheet.
- I will give you a form to sign to show that you said yes to taking part.


What happens if I decide to participate?

- You will be given a questionnaire to complete.

- Another person or I will do an assessment for you.
- The assessment will be to ask you some questions.

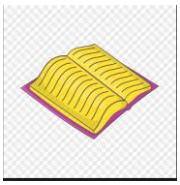
Will I receive treatment for taking part?

- Because it is a research, no treatment will be given.
- We can send you to see your doctor or someone else that can help you.

Who will know about this?

- I will not discuss you with anyone.
- Only my supervisors and I will see your answers.
- Your answers will be kept in a safe place.


- We will write about what we find in a magazine. We will not use your name. You taking part is kept secret.



What happens if I do not like something that's happened?

- I will talk to you to try to help you.
- I will also talk to the person who looks after you.
- If you are unhappy, you or the person who looks after you can fill out a complaint form.
- or call Liz Lukehurst, the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373.


- Or they can help you email my supervisors, Prof Peter Langdon by email Langdon, Peter Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk.

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Appendix 9 – Information sheet and assent form for participants with intellectual disability below 18 years

Participants below 18-years who have ID

TIZARD
University of Kent

Information Sheet and Assent form

Information for participants below 18-years

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents

I am Eziafakaku Nwokolo, a PhD research student.

Purpose of the study

- I want to tell you about some research I am doing.
- Research is a way to learn more about something.
- I would like to find out more about intellectual disability.
- Some children with intellectual disability find it hard to do schoolwork.

Why am I asking you to take part?

- I'm asking you to take part in my research because you have intellectual disability.
- I have spoken to your doctor or your parents or the person that takes care of you at home or in the boarding house you live in.
- They said they are happy for you to take part.

Must I participate?

- You do not have to take part. It is up to you.
- You can say yes now, and you can also say no.
- If you say yes, you can change your mind later.
- If you want to stop, all you have to do is tell me you want to stop.
- I will not be upset at you if you don't want to be in the research.

- I will also talk to your parents, doctor, carer about this research.
- You can talk to them about it before you decide.
- If you want to take part, you will keep this information sheet.
- I will ask you to sign a form to show that you said yes to taking part.

What happens if I decide to participate?

- I will give you some questions to answer.

- The first one will take about 20 minutes for you to finish.

- When you finish, I will mark you answers.

- After marking your answers, one of the people I am working with will ask you more questions.
- It will take about 1 hour.

Will you make me better when I take part?

- Taking part in this study should not affect how you do your school work or homework.
- But, if you get upset, I can talk to your parents or the person who takes care of you.



Is this all kept secret?

- Yes.
- I will not discuss you with anyone.
- Only my supervisors and I will see your answers.
- Your answers will be kept in a safe place.


- We will write about what we find in a magazine. We will not use your name. You taking part is kept secret.

What happens if I do not like something that's happened?

- I will talk to you to try to help you.
- I will also talk to your parents or the person who looks after you.
- If you are unhappy, your parents or the person who looks after you can fill out a complaint form.
- or call Liz Lukehurst, the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373.

- Or they can help you email my supervisors, Prof Peter Langdon by email Langdon, Peter Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk.

Thank you for your time.

Yours sincerely,

Eziafakaku Nwokolo

Would you like to be in this research?

_____ Yes, I will be in this research. _____ No, I don't want to do this.

_____	_____	_____
Child's name	Signature of the child	Date

_____	_____	_____
Person who received assent	Signature	Date

Appendix 10 – Consent form for parents of participants below 18-years with autism or intellectual disability

Parents, carers and guardians as participants and for children below 18yrs already known to have ASD or ID as participants

TIZARD
University of Kent

CONSENT FORM for parents, carers and guardians of participants below 18-yrs with autism

Title: **The reliability and validity of screening tools for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents**

Researcher: Eziafakaku Nwokolo
Email and telephone contact: eun5@kent.ac.uk; +447526316452

Supervisors: Prof Peter Langdon - P.E.Langdon@kent.ac.uk and
Prof Glynis Murphy G.H.Murphy@kent.ac.uk

Please tick to confirm

- I confirm that I have read and understood the information sheet attached for the above study ☐
- I understand that mine and my child/ward's participation is voluntary and that I am free to withdraw my participation or her/him at any time without giving any reason. ☐
- My questions have been answered to my satisfaction. ☐
- I understand that there will be no treatment/intervention offered as part of the research. ☐
- I understand that information I give about my child/ward may be shared within the research team and that it will be kept securely. ☐

- I understand that my child/ward's name will not be used in any publications or presentations arising from the research. ☐
- I agree that my child/ward can participate in the above research study. ☐
- I have read and understood the consent form and agree to mine and my child/ward's participation. ☐

Name of Participant:_____ Signature:_____

Date:_____

Appendix 11 – Consent form for participants above 18-years with intellectual disability or autism

Participants 18yrs and above known to have ASD or ID

TIZARD
University of Kent

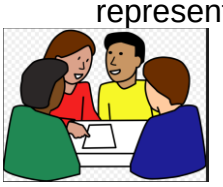
CONSENT FORM for participants 18-years and above

Title: **The reliability and validity of screening tools for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents**

Researcher: Eziafakaku Nwokolo
Email and telephone contact: eun5@kent.ac.uk; +447526316452

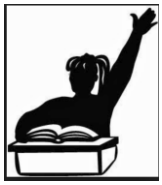
Supervisors: Prof Peter Langdon - Peter.Langdon@warwick.ac.uk and
Prof Glynis Murphy G.H.Murphy@kent.ac.uk

- I have seen the information sheet for the study and talked about it with my representative and the researcher



- I understand that my participation is not by force and that I am free to stop at any time .

- I asked some questions and I was answered properly.



- I understand that I will not be made to feel better by the researchers.

- I know that my answers will be written down and shared with other researchers but they will not know my name.

- I understand that the result may be written in a magazine, but nobody will know my name.

- I would like to take part in the research.

Name of Participant:_____ Signature:_____

Date:_____

Appendix 12 – Information sheet and assent form for participants less than 18-years who have autism

Participants below 18-years who have ASD

TIZARD
University of Kent

Information Sheet and Assent form

Information for participants below 18-years

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents

I am Eziafakaku Nwokolo, a PhD research student.

Purpose of the study

- I want to tell you about some research I am doing.
- Research is a way to learn more about something.
- I would like to find out more about autism.
- Some children with autism find it hard to make friends.
- Some children with autism always like to play with the same toys.

Why am I asking you to take part?

- I'm asking you to take part in my research because you have autism.
- I have spoken to your doctor or your parents or the person that takes care or you at home or in the boarding house you live in.
- They said they are happy for you to take part.

Must I participate?

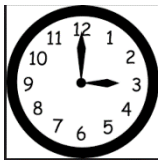
- You do not have to take part. It is up to you.
- You can say yes now, and you can also say no.
- If you say yes, you can change your mind later.
- If you want to stop, all you have to do is tell me you want to stop.
- I will not be upset at you if you don't want to be in the research.

- I will also talk to your parents, doctor, carer about this research.
- You can talk to them about it before you decide.
- If you want to take part, you will keep this information sheet.
- I will ask you to sign a form to show that you said yes to taking part.

What happens if I decide to participate?

- I will ask you some questions and play some games with you.

- It will take about 1 hour for us to finish.



Will you make me better when I take part?

- Taking part in this study should not affect how you are feeling.
- But, if you get upset, I can talk to your parents or the person who takes care of you.



Is this all kept secret?

- Yes.
- I will not discuss you with anyone.
- Only my supervisors and I will see your answers.
- Your answers will be kept in a safe place.


- We will write about what we find in a magazine. We will not use your name. You taking part is kept secret.

What happens if I do not like something that's happened?

- I will talk to you to try to help you.
- I will also talk to your parents or the person who looks after you.
- If you are unhappy, your parents or the person who looks after you can fill out a complaint form.
- or call Liz Lukehurst, the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373.


- Or they can help you email my supervisors, Prof Peter Langdon by email Langdon, Peter Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk.

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Would you like to be in this research?

_____ Yes, I will be in this research. _____ No, I don't want to do this.

_____ _____ _____

Child's name Signature of the child Date

_____ _____ _____

Person who received assent Signature Date

Appendix 13 – Information sheet for those above 18-years with autism

Participants 18yrs and above who have ASD

TIZARD
University of Kent

Information Sheet

Information for participants 18-years and above

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents

I am Eziafakaku Nwokolo, a PhD research student.

Purpose of the study

- I want to tell you about some research I am doing.
- Research is a way to learn more about something.
- I would like to find out more about autism.
- Some people with autism find it hard to make friends.
- Some people with autism always like to do things the same way.
- Some people with autism like to eat the same type of food always.
- Some people with autism always like to talk about the same things.

Why am I being asked to participate?

- I'm asking you to take part in my research because you have autism.
- I have spoken to your doctor or the person that takes care or you at home or in the boarding house you live in.

Must I participate?

- You do not have to take part. It is up to you.
- You can say yes now, and you can also say no.
- If you say yes, you can change your mind later.
- If you want to stop, all you have to do is tell me you want to stop.
- I will not be upset at you if you don't want to be in the research.


- If you want to take part, you will keep this information sheet.
- I will give you a form to sign to show that you said yes to taking part.


What happens if I decide to participate?

- Your parent or guardian or the person that takes care of you will be given a questionnaire to complete.

- Another person or I will do an assessment for you.
- The assessment will be to ask you some questions.

Will I receive treatment for taking part?

- Because it is a research, no treatment will be given.
- We can send you to see your doctor or someone else that can help you.

Who will know about this?

- I will not discuss you with anyone.
- Only my supervisors and I will see your answers.
- Your answers will be kept in a safe place.


- We will write about what we find in a magazine. We will not use your name. You taking part is kept secret.

What happens if I do not like something that's happened?

- I will talk to you to try to help you.
- I will also talk to the person who looks after you.
- If you are unhappy, you or the person who looks after you can fill out a complaint form.
- or call Liz Lukehurst, the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373.


- Or they can help you email my supervisors, Prof Peter Langdon by email Langdon, Peter Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk.

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Appendix 14 – Information sheet for parents of participants below or above 18-years with autism

For parents, carers & guardians of children already known to have ASD

TIZARD
University of Kent

Information Sheet

Information for parents, carers, and guardians of participants below or above 18-yrs with Autism.

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for Intellectual Disabilities and Autism Spectrum Disorder (ASD) in Nigerian adolescents

I am Eziafakaku Nwokolo, a postgraduate student in Intellectual Disabilities at the University of Kent, UK. As part of fulfilling my course requirements, I am conducting a research on the above topic and would appreciate your voluntary participation. For you to do so, some background information has been provided below.

Purpose of the study

This research project is aimed at validating a screening tool for ASD in adolescents. Currently, there are no such tools in Nigeria. The process will involve you completing a questionnaire. This should take between 10 – 20 minutes to complete. We hope that if the research shows that the screening tool is useful, it will become available for doctors, teachers and other health care professionals to use.

What is autism?

Autism spectrum disorder covers a wide range of disabilities that affect social skills, communication and marked by repetitive behaviour. In simple terms, a child or person with autism has trouble with communication, understanding how other people think and may feel and does not like change.

Why have you been given this sheet?

I would like you to participate in this research by responding to some questions concerning your child/ward.

Must I participate?

It is completely voluntary. If you decide to participate, then you will keep this information sheet and be given a consent form to sign. You are free to withdraw your participation at any time without giving a reason.

What happens if I participate?

You will be given a questionnaire to complete. Completion should not take longer than 10 minutes. I will be available to help if you need it. Upon completion, the autism diagnostic observation schedule (ADOS-2) will be used with your child/ward. This is a diagnostic tool that will enable comparison of your responses on the questionnaire with the result of the diagnostic tool. The diagnosis will be done by me or another professional trained to do so.

Will my child/ward receive treatment or intervention as a result of participating?

This is a research project and no treatment will be given.

What about mine and my child/ward's privacy?

Only my supervisors and I will see the returned questionnaires and results, and all personal information will be anonymised. Filled forms will be scanned and saved in an encrypted folder on the computer and hard copies of the questionnaires will be kept in a locked safe.

What happens if either your child/ward or yourself does not like something that's happened?

You can discuss with myself however, you will be given a form at the end of yourself and your child/ward's participation, for feedback or making a complaint. In addition, at any time, you can contact my supervisors, Prof Peter Langdon by email Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk. Alternatively, you can contact the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373. E.Lukehurst@kent.ac.uk

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Appendix 15 – Information sheet for those above 18-years without autism

Participants above 18yrs without ASD

TIZARD
University of Kent

Information Sheet

Information for participants 18-years and above <u>without</u> autism spectrum disorder (ASD)

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for ID and ASD in Nigerian adolescents

I am Eziafakaku Nwokolo, a postgraduate student in Intellectual Disabilities at the University of Kent, UK. As part of fulfilling my course requirements, I am conducting a research on the above topic and I would like to ask you to take part in my project.

Purpose of the study

This project is about validating screening tools for autism in adolescents. Currently, there are no such tools in Nigeria. The process will involve your parent, guardian or someone who knows you well completing a questionnaire. The second part will be carrying out a diagnosis on you using the autism diagnostic observation schedule (ADOS-2). We hope that if the research shows that the screening tools are useful, they will become available for doctors, teachers and other health care professionals to use.

What is autism?

People with autism have difficulties with social skills, communication and repetitive behaviour. This means that they may find it hard forming friendships.

Why am I being asked to participate?

- We need to include people with and without autism in the study for comparison.
- We are asking you to take part because you are someone who does not have autism.

Must I participate?

It is completely up to you. If you decide to take part, then you will keep this information sheet and I'll ask you to sign a form to show that you are happy to take part. You are free to withdraw your participation at any time without giving a reason.

What happens if I decide to participate?

- Your parent, guardian or someone who knows you well will be given a questionnaire to complete. I will be available to help if needed.
- You will be assessed using the ADOS. The diagnosis will be done by me or another professional trained to do so.

What will happen if I receive a positive diagnosis?

This is a research project and no treatment will be given. However, if as a result of the screening you are identified as having autism,

- You will be given information on autism and resources available locally.
- You will be advised to see your health care provider.

What about my privacy?

Only my supervisors and I will see the returned questionnaires, and all personal information will be anonymised. Filled forms will be scanned and saved in an encrypted folder on the computer and hard copies of the questionnaires will be kept in a locked safe.

What happens if I do not like something that's happened?

You will be given a form at the end of your involvement, so that you can give feedback or make a complaint. In addition, at any time, you can contact my supervisors, Prof Peter Langdon by email Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk. Alternatively, you can contact the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373. E.Lukehurst@kent.ac.uk

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Appendix 16 – Information sheet for participants above 18-years without intellectual disability

Participants above 18yrs *without* ID and who are achieving academically at expected levels

TIZARD
University of Kent

Information Sheet

Information for participants 18-years and above without intellectual disability (ID)

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for ID and ASD in Nigerian adolescents

I am Eziafakaku Nwokolo, a postgraduate student in Intellectual Disabilities at the University of Kent, UK. As part of fulfilling my course requirements, I am conducting a research on the above topic and I would like to ask you to take part in my project.

Purpose of the study

This project is about validating screening tools for intellectual disability in adolescents. Currently, there are no such tools in Nigeria. The process will involve two parts (a) you will respond to some questions on a questionnaire and, (b) an intelligence quotient (IQ) test will be administered. The questionnaire should take between 10 – 20 minutes to complete while the IQ-test will take about 1 hour. We hope that if the research shows that the screening tools are useful, they will become available for doctors, teachers and other health care professionals to use.

What is intellectual disability?

People with intellectual disability find it hard to understand new or complicated information and to learn and apply new skills This means they may struggle to look after themselves without help.

Why am I being asked to participate?

- We need to include people with and without intellectual disability in the study for comparison.
- We are asking you to take part because you are someone who does not have intellectual disability.

Must I participate?

It is completely up to you. If you decide to take part, then you will keep this information sheet and I'll ask you to sign a form to show that you are happy to take part. You are free to withdraw your participation at any time without giving a reason.

What happens if I decide to participate?

You will be given a questionnaire to complete. I will be available to help if you need it. Upon completion, a second test for intelligence quotient will be given to you. Your scores from the two tests will be compare. The intelligence quotient test will be done by somebody else on the team.

Will I receive treatment or intervention as a result of participating?

This is a research project and no treatment will be given. However, if as a result of the screening you are identified as having intellectual disability,

- You will be given information on intellectual disability and resources available locally.
- You will be advised to see your health care provider.

What about my privacy?

Only my supervisors and I will see the returned questionnaires, and all personal information will be anonymised. Filled forms will be scanned and saved in an encrypted folder on the computer and hard copies of the questionnaires will be kept in a locked safe.

What happens if I do not like something that's happened?

You will be given a form at the end of your involvement, so that you can give feedback or make a complaint. In addition, at any time, you can contact my supervisors, Prof Peter Langdon by email Peter.Langdon@warwick.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk. Alternatively, you can contact the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373. E.Lukehurst@kent.ac.uk

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Appendix 17 – Information sheet for parents of participants without autism or ID

TIZARD
University of Kent

For parents as participants of children *without* ASD/ID and who are achieving academically at expected levels

Information Sheet

Information for parents, carers or guardians of participants below or above 18-years *without* autism spectrum disorder or intellectual disability

Eziafakaku Nwokolo, Prof Peter Langdon and Prof Glynis Murphy

Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR

Screening for Intellectual Disabilities (ID) and Autism Spectrum Disorder (ASD) in Nigerian adolescents

I am Eziafakaku Nwokolo, a postgraduate student in Intellectual Disabilities at the University of Kent, UK. As part of fulfilling my course requirements, I am conducting a research on the above topic and would appreciate your voluntary participation. For you to do so, some background information has been provided below.

Purpose of the study

This research project is aimed at validating a screening tool for intellectual disability and another for autism spectrum disorder in adolescents. Currently, there are no such tools in Nigeria. The process will involve you completing a questionnaire while your child/ward is administered the diagnostic tool. Completing the questionnaire should take between 10 – 20 minutes to complete. The diagnosis will take about 1-hour. We hope that if the research shows that the screening tools are useful, they will become available for doctors, teachers and other health care professionals to use.

What are intellectual disability and autism?

Intellectual disability simply put means, a serious reduction in a person's ability to understand new or complicated information and to learn and apply new skills (impaired intelligence). The result is a decreased ability to look after oneself independently (impaired social functioning) and begins before adulthood. Autism spectrum disorder covers a wide range of disabilities that affect social skills, communication and marked by repetitive behaviour. In simple terms, a child or person with autism has trouble with communication, understanding how other people think and may feel and does not like change.

Why have you been given this sheet?

- I would like you and your child/ward to participate in the study.
- I am asking you either as a parent, carer or guardian of the individual who does not have ID or ASD and is meeting set targets, performing at similar levels as his/her peers and does not engage in any behaviour that gives you concern, to complete the questionnaire.
- I am asking for your consent as the parent, carer or guardian of a minor to consent on their behalf for the diagnosis to be done.

Must either my child/ward or myself participate?

It is completely voluntary. If you decide to participate and allow your child/ward to participate, then you will keep this information sheet and be given a consent form to sign. Your child who is under 18-years will be given an assent form to sign. You are free to withdraw yours or your child/ward's participation at any time without giving a reason.

What happens if we participate?

- You will be given a questionnaire to complete.
- For ID, your child/ward will be given a questionnaire to complete. Completion should not take longer than 10 minutes. I will be available to help if s/he needs it.
- Upon completion of the questionnaires by both of you, further diagnosis will be done to compare the scores with the result of the diagnostic tool. The diagnosis will be done by me or another professional trained to do so.

What about our privacy?

Only my supervisors and I will see the returned questionnaires, and all personal information will be anonymised. Filled forms will be scanned and saved in an encrypted folder on the computer and hard copies of the questionnaires will be kept in a locked safe.

What happens if either your child/ward or yourself does not like something that's happened?

You can discuss with myself however, you will be given a form at the end of your participation, for feedback or make making a complaint. In addition, at any time, you can contact my supervisors, Prof Peter Langdon by email P.E.Langdon@kent.ac.uk or Prof Glynis Murphy G.H.Murphy@kent.ac.uk. Alternatively, you can contact the Secretary of the Ethics Committee, Tizard Centre, University of Kent, Canterbury, Kent, CT2 7LR, Telephone: +44(0)1227 827373. E.Lukehurst@kent.ac.uk

Thank you for your time.
Yours sincerely,
Eziafakaku Nwokolo

Appendix 18 – Consent form for participants above 18-years without an intellectual disability or autism

Participants 18-yrs and above without ASD/ID and who are achieving academically at expected levels

TIZARD
University of Kent

CONSENT FORM for participants 18-years and above.

Title: **The reliability and validity of screening tools for Intellectual Disabilities and Autism Spectrum Disorder in Nigerian adolescents**

Researcher: Eziafakaku Nwokolo
Email and telephone contact: eun5@kent.ac.uk; +447526316452

Supervisors: Prof Peter Langdon - Peter.Langdon@warwick.ac.uk and
Prof Glynis Murphy G.H.Murphy@kent.ac.uk

Please tick to confirm

- I confirm that I have read and understood the information sheet attached for the above study ☐
- I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. ☐
- My questions have been answered to my satisfaction. ☐
- I understand that there will be no treatment/intervention offered as part of the research. ☐
- I understand that information I give may be shared within the research team and that it will be kept securely. ☐
- I understand that my name will not be used in any publications or presentations arising from the research. ☐

- I agree to take part in the above research study. ☐
- I have read and understood the consent form and agree to participate. ☐

Name of Participant:_____ Signature:_____

Date:_____

Appendix 19 – Participants feedback form

Feedback / Comments Form

Thank you for agreeing to talk to Eziafakaku Nwokolo to help with her research about 'The reliability and validity of screening tools for and Autism Spectrum Disorders and Intellectual and Developmental Disabilities in Nigerian adolescents'.

We hope that everything was alright when you talked to Eziafakaku Nwokolo. We would be interested in any comments you would like to make, positive or negative.

When things go well, we like to encourage researchers by giving them good feedback. But if things don't go well, it will help us to know this.

Please send any comments you have to:

Liz Lukehurst

Secretary to the Tizard Centre Research Ethics Committee

Tizard Centre

Cornwallis North East

University of Kent

CT2 7NF

Or E.Lukehurst@kent.ac.uk

Thank you once again for helping the Tizard Centre with our research.

Tizard Centre Research Ethics Committee

Appendix 20 – Systematic Review Published Manuscript

Link to the article - https://doi.org/10.1007/s40489-022-00342-6

Screening for intellectual disabilities and/or autism amongst older children and young adults: A systematic review of tools for use in Africa

Abstract

There are many well-developed screening tools for both intellectual disabilities and autism, but they may not be culturally appropriate for use within Africa. Our specific aims were to complete a systematic review to: (1) describe and critically appraise short screening tools for the detection of intellectual disabilities and autism for older children and young adults, (2) consider the psychometric properties of these tools, and (3) judge the cultural appropriateness of these tools for use within Africa. Six screening tools for intellectual disabilities and twelve for autism were identified and appraised using the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines. We identified two screening tools which appeared appropriate for validation for use within African nations.

Keywords: adolescent, screening, diagnosis, autism, intellectual disabilities, Africa

Several studies (e.g., Eldevik et al., 2009; Steiner et al., 2012; Swinkels et al., 2006; Luckasson & Schalock, 2013; Schalock & Luckasson, 2013) have highlighted the benefits of early detection of developmental disabilities such as intellectual disabilities and autism. The benefits have included improved behavioural outcomes and family support, as well as earlier intervention. Other benefits included improved planning for educational needs and support, improved social skills, and greater cognitive and language development. These findings have emerged predominantly from Western and high-income countries with there having been very limited research from low to medium-income countries (LMICs), as indexed by the published gross national income by the United Nations (Tomlinson et al., 2014; Gladstone et al., 2010; United Nations, 2014; World Bank, 2020 & United Nations Department of Economic and Social Affairs, 2021). While the presentation of autism is the same regardless of economic status, the political climate and associated social burdens within LMICs, such as in the African countries, discourages the early detection of developmental disabilities as it is not seen as urgent, which increased the health disparities faced by this population (Emerson, 2012; Gladstone et al., 2014).” The situation is similar for those with intellectual disabilities, with late identification leading to further delay of intervention.

Screening for developmental disabilities can be done in any setting, such as the community (Kopp & Gillberg, 2011), schools (Suhail & Zafar, 2008; Webb et al., 2003), primary care settings (Robins, 2008; Barton et al., 2012; Gura et al., 2011; Limbos et al., 2011), urban settings (Guevara et al., 2013), the criminal justice system (Murphy et al., 2017), and many others. In the African context, individuals with developmental disabilities are noticed either in schools, or when parents seek medical attention for a severe illness, or when researchers embark on studies targeted specifically at populations with disabilities (Knox et al., 2018; Saloojee et al., 2007; Gladstone et al., 2010; Scherzer et al., 2012).

Preliminary screening for intellectual disabilities or autism can occur through the use of a variety of methods, such as observation, informal and formal interviews, history taking and the use of short screening tools. Irrespective of which method is used, the important factors to consider are the accuracy of results, validity, reliability, training requirements, ease of administration and the simplicity and ease of interpreting results (Westerlund & Sundelin, 2000; Cochrane & Holland, 1971). The accuracy of screening tools is vital, and Glascoe (2005) recommends that the

sensitivity, or the true positive rate, should be between 70-80%, while specificity, or the true negative rate, should be at least 80%. Screening tools require validation when used outside the environment and population for which they were developed, and this process involves comparing the results of the screening tool to that of an accepted gold standard instrument (Maxim et al., 2014). Most screening tools in existence have been validated in the West, but evidence for their validation in Africa is scant (Soto et al., 2015; Van der Linde et al., 2015).

Another consideration is the adaptation of measures for use outside of the original design environment. A robust screening tool should be culturally sensitive and useable with multiple populations (Van der Linde et al., 2015). Given that almost all the measures were developed within Western countries, issues regarding cultural sensitivity and feasibility of using these screening tools in their original format with the African populace needs investigating. Screening tools developed in high-income environments do not necessarily consider the application and understanding of the terminology in other environments. Screening results and reliability can be affected where the language of the screening tool differs in application or understanding (Soto et al., 2015). In Africa, some studies that measured developmental milestones and disabilities utilised screening tools developed in the West (Oshodi et al., 2016; Koura et al., 2013; Jinabhai et al., 2004). For example, Oshodi et al. (2016) used the Modified Checklist for Autism in Toddlers (M-CHAT) in a Nigerian urban setting where language and terminology were not barriers, thereby eliminating the need for translation. Jinabhai et al. (2004) adapted and substituted examples in both the Auditory Verbal Learning Test (AVLT) and Young's Group Mathematics Test (GMT) with more familiar items for Zulu participants. The AVLT instructions were given in Zulu with the items 'turkey' and 'ranger' replaced with the more culturally familiar Zulu words 'chicken' and 'herdboy' respectively. Jinabhai et al. (2004) made considerably more adaptations to the GMT and administered the test in Zulu. The adaptations centred on change of words and examples to more familiar items such as 'tarts' to 'cakes', 'marbles' to 'balls', 'engine' to 'truck' and the names 'Dick and Jim' were changed to 'Sipho and Thembi'. Koura et al. (2013) adopted a rigorous translation model to translate and adapt the Mullen Scales of Early Learning (MSEL) used in their study from English to French while the parents' instruction was translated into Fon, the local language.

There are some other studies which highlight the importance of language and terminology in translated versions. Wild et al. (2012) translated the CBCL into six languages Korean, Hebrew, Spanish, Kannada, and Malayalam. In the Malayalam version, several cultural adaptations such as changing the "milk delivery" to a more familiar job and giving different examples of sports and hobbies were needed, while in the Hebrew version two sexually related items were removed from the measure. Koura et al. (2013) also used the "Ten Questions" (TQ) to screen for disabilities and collect cognitive development information for their participants; however, the items were not translated into other languages.

The "Ten Questions" is a disability screening tool which has been used widely in developing countries. The TQ was primarily designed as a stop-gap screening tool for numerous kinds of impairment in children aged 2 – 9 years old, including intellectual disabilities, and has been used to estimate prevalence within low-income and low-resource countries. The TQ is about cognitive skills, motor skills, hearing, epilepsy, and vision problems. Stein et al. (1986) used the measure as a screening tool in the first stage of their prevalence study across several countries, to identify children with moderate to profound intellectual disabilities. Intellectual disabilities were classified as an intelligence quotient less than or equal to 55 ($IQ \leq 55$). Study samples were from eight countries (India, Philippines, Bangladesh, Sri Lanka, Malaysia, Pakistan, Brazil, and Zambia). No specific figures were reported for sensitivity and specificity. However, the team reported that most participants with intellectual disabilities were probably identified while children with other conditions and IQs greater than 55, were also identified. These results from Stein et al. (1986) do not seem adequate to judge the psychometric properties of the TQ. Also, any consideration of cultural issues was not documented. Two other studies, Mung'ala-Odera et al., (2004) in Kenya and Kakooza-Mwesige et al. (2014) in Uganda, used the TQ with children in their early years. Kakooza-Mwesige and colleagues screened 1,169 Ugandan children between the ages of 2 and 9 years using an adapted version of the TQ, which included 13 additional questions about autism. Questions about autism covered the three criteria: qualitative impairment in social interaction, qualitative impairments in communication and restricted repetitive and stereotypical behaviours. The adapted version of the TQ was called the 23Q. The authors reported high negative predictive value (.90) and specificity (.90) with very low positive predictive value (.22) and sensitivity (.52) for participants with autism. As such Kakooza-Mwesige et al. (2014) concluded that the neither the TQ nor the 23Q met the criteria as useful

screening tools for autism. While the TQ has been useful in identifying children with specific disabilities, its appropriateness for detecting more complex and hidden disabilities such as mild intellectual disabilities or autism are unclear (Olusanya & Okolo, 2006; Durkin, 2001). Also, suggestions have been made that the continued use of the TQ in Africa or LMICs may undermine efforts towards effective screening and early intervention (Olusanya & Okolo, 2006).

Additionally, screening tools are sometimes used to monitor the progress of interventions, where the same measure is re-administered to the same individual, to examine progress, indicative of the "responsiveness" of a screening tool. McConachie et al. (2015) reviewed the measurement properties of some screening tools used to measure progress and outcomes in young children with autism spectrum disorder aged up to 6 years. Their reviewed focused on measuring the progress and improved quality of life post intervention for participants in the West.

Soto et al. (2015), in their systematic review of 21 included studies, investigated efforts towards the cultural adaptation of screening tools for use outside of the environments in which they were primarily developed. With a specific emphasis on autism spectrum disorder only, the review examined the adherence to recommended adaptation procedures and the psychometric properties of the adapted instruments. Studies about people with intellectual disabilities were excluded. The adaptation studies included in the Soto et al. review had been carried out in nineteen countries and involved ten languages. Only two of those countries are in continental Africa: Egypt, and Tunisia. The M-CHAT was used in the studies in both countries. Egypt and Tunisia are Arabic-speaking countries and the M-CHAT was translated into Arabic. In LMICs, where resources are limited, the cost and burden of a rigorous translation and adaptation process is a barrier to acquiring reliable screening tools.

Recently, attempts have been made towards developing screening tools in areas such as nutrition, neurodevelopmental disabilities and mental health which are culturally sensitive for use within the African continent (Gladstone et al., 2010; Hasegawa et al., 2017; Vawda et al., 2017). While these efforts are commendable, study populations are often limited to early childhood, with children aged 2- to 9-year-olds. The focus upon young children (2- to 5-year-olds) would allow for the implementation of interventions earlier, but would miss older children (10 years and above). Relative to studies on young children, there is very little data on studies with older children and adolescents; however, studies involving adolescents are emerging (Allison et al.,

2012; Morales-Hidalgo et al., 2017; Nijman et al., 2016). The paucity of adolescent studies is not peculiar to Africa. What appears to be unique to LMICs and Africa is the relatively low level of awareness, insufficient economic resources, insufficient numbers of professionals, and a culture of not seeking immediate help (Franz et al., 2017).

In African countries like Nigeria, Kenya, Ghana, and Uganda, awareness is growing, yet it is still common for families not to seek immediate help for individuals with autism or intellectual disability till later in life (Franz et al., 2017). It therefore remains the case that many of these children are not screened or diagnosed early in life. Such individuals are then brought to the attention of professionals around the onset of adolescence as this is the period when teenagers begin to spend an increasing amount of time away from the family home. Adolescents and young people are those aged 11 to 26 years of age, an age range which is consistent with the critical period of brain maturation associated with development during adolescence (Sawyer et al., 2012; 2018). To identify these older children and young adults who have been missed or not diagnosed in a time-efficient and effective way, an appropriate screening tool should be available. However, there is a marked absence of well-developed screening tools for use with adolescents among professionals and services in African countries (Hirota et al., 2018).

Overall, screening for either intellectual disabilities or autism in individuals in African countries requires the use of a validated and reliable measure which is accessible to front line professionals such as teachers, nurses, carers, family doctors and those who are in primary health care services. While some screening tools have been developed and validated in the West, and investigated for use in Africa, the researchers have not always compared their study results against acceptable gold standard instruments, a crucial stage in measuring the validity of tools when used in new environments. For instance, Oshodi et al. (2016) and Koura et al. (2013) obtained reasonable results from their studies. However, they did not compare their results to that of an acceptable gold standard instrument and this presents limitations. Besides selecting and validating a standardised screening instrument for use with adolescents, the tool ought to be culturally relevant for use within the African context. Through careful adaptation and translational work, screening tools developed in the West may be adopted for use in LMIC such as Nigeria, Ghana and other African countries. By doing so, some of the costs and time to develop entirely new tools can be reduced.

To identify such tools, a systematic review was completed with the following aims: (1) to describe and critically appraise short screening tools for the detection of intellectual disabilities and autism in children and young people aged 11 to 26 years, (2) to consider the psychometric properties of these tools, and (3) to consider the appropriateness of using these tools across a range of cultures.

Method

Search Strategy

A literature search of the following electronic databases was carried out to identify relevant studies: Academic Search Complete, MEDLINE, CINAHL Plus, PsycINFO and PsycArticles. The key search terms were 'intellectual', 'learning', and 'autism'. These key terms were then combined with disability and with screening and diagnosis. Truncated terms were used as appropriate to ensure inclusion of variations of the words. Older words used to describe people with intellectual disabilities, such as 'mentally retarded' or 'mental retardation' were also included. Titles and abstracts were the focus of the initial search. The combined search terms are found in Table 1. Backward (ancestry) searching was used to identify other papers that may be relevant from references of eligible studies. The search was done using EBSCOhost and concluded on the 22nd of June 2018. To ensure that no new studies published, or tools developed were missed, the search was updated with the same terms on the 5th of November 2020.

Table 1 About Here

To provide transparency of the review process and avoid duplication of the study, the review was registered with Research Registry (https://www.researchregistry.com/ - Registration Code: reviewregistry798). Research Registry is an international database for registering all types of research studies such as case reports, observational and interventional studies, systematic reviews and meta-analyses.

Eligibility Criteria and Study Selection

Titles and abstracts were initially screened for inclusion based on the following criteria: (1) the article was written in English, (2) validated screening tools were used, or the study involved developing a screening tool, (3) little or no extra training was required to administer the tool, (4) the tool did not take longer than 1-hour to administer, (5) some or the majority of the participants were aged 11 years and younger than 27 years, and (6) participants in the validation sample for intellectual disabilities or autism were diagnosed by a duly qualified healthcare professional. Some articles which had multiple studies and participants across a broad age range (Baron-Cohen et al., 2006; McKenzie et al., 2012c; Nijman et al., 2016; Deb et al., 2009; Kraijer & De Bildt, 2005) were included because of their relevance in at least part of their research. Studies were excluded if any of the following criteria were met: (1) the study was related to other health issues such as diabetes, cancer, visual and any other medical condition in persons with intellectual disabilities or autism, (2) the study was about linguistic and speech-related conditions, (3) the study was about developmental learning disorders/difficulties (e.g. impairments in reading or writing) (4) the tools were not for screening but diagnostic tools, (5) publications were letters, correspondences, editorials or recommendations to the editors, (6) studies had missing information on age, (7) full text was not available, and (8) additional skills or training were required to administer the tool. Due to the paucity of research with adolescents, and in order not to miss any potential screening tools, there was no restriction on publication date. Studies done in both clinical and non-clinical settings were considered. Also, the inclusion of English only articles was based on the authors language proficiency. This initial search produced over 1000 potential articles.

After removing duplicates and completing a title and abstract screen against the eligibility criteria, a total of 235 articles were retrieved for full-text screening. This led to the exclusion of a further 194 papers. Studies were excluded due to the ages of participants (n=70) or the fact that the article was not about screening tools (n=48). One of the papers excluded at this stage had no participants in the study (Al Mamun et al., 2016), another had no details of the author and the full text could not be accessed, and thirty-three were about specific learning disorders/difficulties. See Figure 1 for the PRISMA flowchart of the study selection process (Page et al., 2021). The remaining 41 studies met eligibility criteria. The eligibility criteria were

applied independently by two members of the research team (EN & GM) with excellent agreement, $k = 1$.

Figure 1 About Here

Quality Appraisal

A quality appraisal of the included studies was conducted using the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) risk of bias checklist (Terwee et al., 2018; Prinsen et al., 2018 & Mokkink et al., 2018a) and the manual as guides. The COSMIN manual was developed for the systematic review of patient-reported outcome measures (PROMs). Although this review did not consider PROMs, the COSMIN checklist was adopted due to its robustness. The relevant aspects of the COSMIN for this review include its usefulness for assessing the methodological quality of studies, development and design of measurement tools, psychometric properties, and cultural validity. The appraisal was done for all papers by EN and was independently checked by a second member of the team (PL) for 40% of the papers. Following the review of the ratings and correction of errors, the agreement was $k = 1$. Based on the COSMIN guidelines, the quality of included studies was rated (Tables 2 and 3 in Appendix). For each study, the quality was assessed based on a four-point rating system where each standard within the COSMIN box can be rated as ‘very good’, ‘adequate’, ‘doubtful’ or ‘inadequate’. This overall rating of the study quality contributed to grading the quality of the evidence for each tool. The quality of evidence and methods were scored on a four-point rating scale, that is, sufficient, insufficient, indeterminate, or inconsistent. The overall score for quality of evidence is according to the “lowest score counts” method, and the categories used were high, moderate, low, and very low. Overall ratings for the study methodologies, quality of tool development and quality of evidence for the measurement properties using the COSMIN checklist are in Tables 4 to 7 (in Appendix).

One key component of the COSMIN is its usefulness in evaluating cross-cultural validity of tools. Cross-cultural validity refers to “the degree to which the performance of the items on a translated or culturally adapted PROM (Patient Reported Outcome Measures) are an adequate

reflection of the performance of the items of the original version of the PROM" (Prinsen et al., 2018, p. 1154). Cross-cultural validity is assessed when a tool is used with at least two different groups. Such populations could differ in language, diagnosis, gender, age groups or ethnicity (Mokkink et al., 2018b).

Additionally, the COSMIN manual suggests areas for adaptation by the review team. Some of the adaptations made for this review were related to the hypothesis testing for responsiveness (criterion validity) and construct validity. In the case of Box 9 which is hypothesis testing for construct validity, we used it to assess the convergent validity and discriminative validity where applicable. Regarding responsiveness, Box 10a for criterion approach, we assessed the diagnostic accuracy of diagnostic tools used in the studies rather than change scores. Outcome measures of specificity and sensitivity were also assessed. For Boxes 10b and c, construct approach, studies which utilised similar measurement instruments or where the study design was between groups (children, adults, or those with and without ID or ASD sub-groups). Box 10d was not utilised for any studies as we did not look at interventions. Ratings of insufficient, inadequate, or doubtful were given in instances where there was insufficient information reported in the study for a higher rating as required by the COSMIN checklist. For clarification and completeness, manuals, where available, and authors of the tools were consulted for further evidence. This is discussed further in the result section.

Data Extraction and Synthesis

Relevant information about the aims of each included study, along with the tool used, the design, participants, time to administer the tool, and outcomes were extracted and are reported in Tables 8 and 9 (in Appendix). The tables were arranged alphabetically by the first author, and chronologically when the first author co-authored more than one study. All included studies were quantitative.

Results

Search Results

Forty-one papers met the eligibility criteria (Figure 1), and 22 of these were about screening tools for autism, while 19 focused upon screening tools for intellectual disabilities. The quality ratings for the included studies are found in Tables 2 and 3 (in the Appendix). Additionally,

sensitivity (true positive rate), specificity (true negative rate), positive predictive value or precision (the probability of screening positive and being correct), and negative predictive values (the probability of screening negative and being correct) for the tools were extracted and are reported in Tables 8 and 9.

Description and Characteristics of Studies

Autism Spectrum Disorder

There were a total of 12,240 participants across the 22 autism studies with age ranging from 15-months to 80-years. Studies with children younger than 11-years or older than 27-years old were only included if most of their participants were within the specified range of the inclusion criteria. Of the 12,240 participants, a little over 9,000 involved proxy respondents such parents, teachers, or caregivers of people with autism. Most of the studies were conducted in the United Kingdom (UK) (n = 7) with others spread across various countries, including the United States (US) (n = 5), Spain (n = 2) and one each from Germany, Sweden, Netherlands, Qatar, Australia, Turkey, Singapore and Argentina (Table 8). A variety of screening tools were used across the studies, including the Autism Screening Quotient (AQ-10) adolescent and adult versions (n = 3), Autism Screening Quotient (AQ-50) (n = 1), Autism Spectrum Screening Questionnaire (ASSQ) (n = 1), Autism Spectrum Screening Questionnaire-Revised Extended Version (ASSQ-REV) (n = 1), Social Communication Questionnaire (SCQ) (n = 7), Developmental Behavior Checklist-Autism Screening Algorithm (DBC-ASA) (n = 1), Childhood Autism Rating Scale (CARS) (n = 1), Mobile Autism Risk Assessment (MARA) (n= 1), Pervasive Developmental Disorder in Mentally Retarded Persons Scale (PDD-MRS) (n = 2), EDUTEA (n = 1), Child Behavior Checklist (CBCL) (n = 1), Adapted Autism Behaviour Checklist (AABC) (n = 1), and Autism Diagnostic Inventory-Telephone Screening in Spanish (ADI-TSS) (n =1).

Fifteen studies were between-group designs, one within-group design, and six single group designs. One study was longitudinal and included data collected over 15-years. Across the included 22 studies, two broad aims were discerned: (a) designing a short screening tool, and (b) validating the discriminative ability of tools. Those that focused on designing short tools were further categorised in two ways: (a) adapting existing tools into shorter versions, or (b) the development of entirely new tools. Eighteen out of 22 (approximately 82%) of the papers reviewed based their studies on existing tools developed over ten years ago. The remaining four

(18%) considered tools that were developed in the last two to three years. The existing tools were mainly used with children, while the other studies reviewed focused upon adapting the tools for adolescents and adults.

Autism Screening Tools

The Autism Spectrum Quotient (AQ). The Autism Spectrum Quotient (Baron-Cohen et al. 2001) is a short, easy to use and score, self-administered screener for adults with Asperger Syndrome or High Functioning Autism. It is comprised of 50 questions divided into five subsets of 10 questions each covering five domains – social skills, attention to detail, attention switching, communication and imagination. Over time, the AQ was adapted and modified to include adolescents (Baron-Cohen et al., 2006) while maintaining the original 50-item format. The AQ-50 child, AQ-50 adolescent and AQ-50 adult measures were adapted to create shorter versions by selecting the ten most discriminating items from each and validating the short tool (Allison et al., 2012). The AQ in different variations was used in three different studies (Allison et al., 2012; Baron-Cohen et al., 2006; Booth et al., 2013). The short version of the adolescent tool, the AQ-10 (Allison et al., 2012) had a sensitivity of .93, a specificity of .95 and a positive predictive value (PPV) of .86 while the AQ-50 (Baron-Cohen et al., 2006) tested with adolescents had a sensitivity of .89 and specificity of 1. Baron-Cohen et al. (2006) reported no PPV, but commented that future research should explore this.

For adult participants (includes participants older than 18 years and/or 16 years of age in some instances) that employed the short AQ-10, Allison et al. (2012) found a sensitivity of .88 and specificity of .91 while Booth et al. (2013) found a sensitivity of .80 with a specificity of .87. All three studies (Allison et al., 2012; Baron-Cohen et al., 2006; Booth et al., 2013) included participants with a previous diagnosis of autism.

While all the three studies that employed the AQ defined the constructs to be measured, the quality of evidence was rated as low. Specifically, content validity was rated as low since participant and expert involvement in the studies was unclear. Structural validity, internal consistency, reliability, construct validity, cross-cultural validity and criterion validity were all examined by Allison et al. (2012). Baron-Cohen et al. (2006) examined internal consistency and reliability with moderate evidence for cross-cultural validity. Booth et al. (2013) provided evidence for structural validity, while reliability and cross-cultural validity were undetermined.

As such, the evidence for reliability was rated as low and the overall rating for cross-cultural validity was found to be low. To ensure that the psychometric properties of the AQ-10 were accurately captured, the authors were contacted for the manual, who responded that the tests and 'manuals' were those on the authors' website. In summary, although the psychometric results met the criteria for good tools (Glascoe, 2005) following the COSMIN guidelines, where the lowest score counts, the overall quality of evidence for the tool was rated as low.

Autism Spectrum Screening Questionnaire (ASSQ) and the Autism Spectrum Screening Questionnaire-Revised (ASSQ-REV). Preliminary development of the ASSQ took place in Sweden for use within a prevalence study for high-functioning autism and Asperger syndrome in mainstream schools (Ehlers and Gillberg, 1993). The ASSQ is a 27-item checklist that can be completed by laypersons such as teachers or parents and was developed further in later studies (Ehlers et al., 1999). An extended version of the ASSQ-REV was developed for the early identification of girls with autism (Kopp & Gillberg, 2011). The original Swedish version of the ASSQ has been translated into multiple languages – Mandarin Chinese (Guo et al., 2011), English (Ehlers & Gillberg, 1993), Norwegian (Posserud et al., 2006), Finnish (Mattila et al., 2009), and Lithuanian (Lesinskiene, 2000).

Cederberg et al. (2018) examined the diagnostic accuracy of the ASSQs in adolescents previously diagnosed with high functioning autism. While participant gender and the psychometric properties of the measure were not reported, the authors reported that the ASSQ appeared sensitive to correctly identifying autism. Kopp & Gillberg (2011) examined the validity and accuracy of individual items for detecting autism in girls and boys aged 6 – 16yrs. Different items showed considerable discriminative ability (AUC > .70, see Kopp & Gillberg, 2011) for those with autism versus typically developing children across genders. Both studies used participants who had a previous diagnosis of Autism. Like Cederberg et al. (2018), Kopp & Gillberg (2011) reported no sensitivity, specificity, PPV or negative predictive value (NPV).

Although the ASSQ was originally in Swedish, and has been translated into different languages, cross-cultural validity was rated as low using COSMIN due to insufficient evidence of its effectiveness in different cultures. Criterion validity, construct validity, internal consistency and reliability were rated as insufficient based on the combined evidence from both studies (Kopp & Gillberg, 2011; Cederberg et al., 2018). Neither of the two studies examined the content nor

structural validity of the ASSQ. To ensure that all relevant evidence and information on the tool's development were examined, efforts were made to access the manual but were unsuccessful. No other studies utilising the ASSQ outside the West were found. The overall quality of the tool was rated as very low.

Social Communication Questionnaire (SCQ). The SCQ formerly known as the Autism Screening Questionnaire (ASQ; Berument et al., 1999), was initially designed as a companion screening tool for the Autism Diagnostic Interview (ADI) (Snow, 2013). The SCQ is a brief 40-item parent or caregiver-report screening measure modelled after the ADI-R and has been used widely in research (Berument et al., 1999; Rutter et al., 2003). The measure has two versions; the lifetime version and the current version, both focusing on symptoms of autism most likely to be observed by the individual's principal caregiver. The caregiver must be familiar with the individual's developmental history and current behaviour. The SCQ is a screening tool and cannot be used for the diagnosis of autism. The measure is used for anyone 4-years old and above. The design allows for the comparison of symptoms across different groups of individuals such as children with language delays and those with medical conditions co-existing with autism. The SCQ is currently available in seventeen languages (Danish, Dutch, English, Finnish, French, German, Greek, Hebrew, Hungarian, Italian, Japanese, Korean, Norwegian, Romanian, Russian, Spanish and Swedish) and is used widely in research.

Seven studies (Aldosari et al., 2019; Brooks & Benson, 2013; Berument et al., 1999; Charman et al., 2007; Corsello et al., 2007; Mouti et al., 2019; Ung et al., 2016) utilised the SCQ. Five studies (Aldosari et al., 2019; Charman et al., 2007; Corsello et al., 2007; Mouti et al., 2019; Ung et al., 2016) included samples of adolescents, one included adults with intellectual disabilities (Brooks & Benson, 2013), while one (Berument et al., 1999) was a development study and included children, teenagers and adults (age range: 4 – 40 years). Berument and colleagues (1999) recommended an optimal cut-off of 15 for differentiating those with and without autism. Using this cut-off, they reported a sensitivity of .85, specificity of .75, PPV .93 and NPV .55. In the other studies, the cut-off was varied to generate optimal values, depending on the age of the participants. For instance, Brooks & Benson (2013) using a cut-off of 15, reported that the sensitivity was .71, specificity .77, PPV .58 and NPV .86. However, when the cut-off was lowered to 12, the sensitivity was .86, specificity .60, PPV .49 and NPV .91. Similarly, Corsello

et al. (2007) reported finding sensitivity of .71, specificity = .71, PPV = .88, and NPV = .45 at a cut-off of 15 while at a cut-off of 12 sensitivity was .82, specificity = .56, PPV = .84, and NPV = .51. However, as is typical with screening tools, lower cut-off scores will improve sensitivity, but at the expense of specificity.

Recently, Mouti et al. (2019) examined the optimal cut-off for differentiating between ASD, attention deficit and hyperactivity disorders (ADHD) and typically developing individuals. Their result showed that at a cut-off of score of 9, the SCQ showed excellent discriminative ability between ASD and Non-ASD with a sensitivity of .1 and specificity of .84. Additionally, Mouti et al. (2019) showed that at the cut-off of 13, ASD was clearly discriminated in individuals who were diagnosed as ASD only (sensitivity = .96, specificity = .87) or a combination of both ASD and ADHD (sensitivity = .87, specificity = .85). In the Arabic validation study, Aldosari et al. (2019) reported sensitivity and specificity of .80 and .97 respectively at the recommended cut-off score of 15. However, for a cut-off range between 11 and 15, the sensitivity varied between .90 and .80 while specificity varied between .85 and .97. Aldosari et al. (2019) also reported internal consistency of $\alpha = .92$.

Apart from Ung et al. (2016), who validated the SCQ against the Childhood Autism Rating Scale (CARS-2) only, all the other studies validated the SCQ against either the Autism Diagnostic Observation Schedule-2 (ADOS-2) or Autism Diagnostic Interview-Revised (ADI-R) or a combination of the CARS, ADOS-2 and ADI-R. Overall, the psychometric properties of the SCQ met the guidelines (Glascoe, 2005) for good tools, and the SCQ correlated well with the ADI-R (Berument et al., 1999).

Out of the seven studies reviewed, four (Berument et al., 1999; Corsello et al., 2007; Mouti et al., 2019; Aldosari et al., 2019) examined the structural validity with sufficient outcomes reported. Criterion validity and reliability were rated as excellent across all seven studies. All seven studies had clear constructs with five (Charman et al., 2007; Corsello et al., 2007; Ung et al., 2016; Mouti et al., 2019; Aldosari et al., 2019) providing sufficient evidence for the construct validity. There was an excellent outcome on the criterion validity across all seven studies. Five studies (Berument et al., 1999; Corsello et al., 2007; Mouti et al., 2019; Aldosari et al., 2019; Charman et al., 2007) rated positive had sufficient evidence for cross-cultural validity while the remaining two (Brooks & Benson, 2013; Ung et al., 2016) were rated negative with insufficient

evidence. Soto et al. (2015) in their review of culturally adapted tools, reported that the Chinese validation study (Gau et al., 2011) of the SCQ had good test-retest reliability (r_{ICC} = .77 – .78) and internal consistency (α = .73 – .91). The authors (Gau et al., 2011) reported excellent concurrent validity ($r \leq .65$). Given that the SCQ is available in 17 languages, has been used across countries including Africa (Bozalek, 2013), across ethnicities, genders, ages, and widely employed in research, it meets several of the qualities for good cross-cultural validity as defined by COSMIN. The SCQ was rated overall as medium based on the evidence from the seven studies reviewed (Brooks & Benson, 2013; Berument et al., 1999; Charman et al., 2007; Corsello et al., 2007; Ung et al., 2016; Mouti et al., 2019; Aldosari et al., 2019) and previous work done by McConachie et al. (2015). Given the above results, the SCQ seems an appropriate tool to be considered for use within African nations, especially as very little training is required to score it.

Childhood Autism Rating Scale (CARS). The CARS was developed by Schopler et al. (1980) as a diagnostic tool for children with autism. However, this measure, while meant to be diagnostic, was included because Mesibov et al. (1989) used it as a screening tool with adolescent participants, suggesting the CARS' potential utility as a screening instrument for autism. Nevertheless, Mesibov et al. (1989) did comment that the CARS was meant to be used as a diagnostic tool. The CARS is a 15-item rating scale that assesses behaviours associated with autism. The measure is meant to ease the identification of children with autism for parents, educators, clinicians, and other health care providers. The scale is available in English, Brazilian Portuguese, Lebanese, Japanese, Swedish, and French. The second edition now includes a scale for identifying high functioning autism and a parent information form. Some training is required to administer the tool.

Although the CARS was initially validated for use with children, Mesibov et al. (1989), in their longitudinal study, examined its suitability for use with adolescents and adults with autism. Fifty-nine participants with a previous autism diagnosis were re-assessed, and the results showed that 81% (n = 48) retained their diagnosis. In comparison, 19% (n =11) of them received a revised diagnosis of no autism based on a cut-off score of 30. However, moving the score to a cut-off of 27 (to account for the mean difference in scores between the younger and older sample), 92% (n = 54) were accurately diagnosed. As a result of the improved diagnostic

outcomes, Mesibov and colleagues recommended 27 as the cut-off for persons over the age of 13-years.

Based on COSMIN guidelines, content validity, internal consistency, and construct validity were rated as not determined, since it was unclear from the study whether these were tested. There was insufficient evidence for structural validity, and criterion validity. Cross-cultural validity was rated as positive with moderate evidence due to the availability of the measure in different translations. The evidence for reliability was moderate; however, this was based on the evidence from the only study found (Mesibov et al., 1989). Authors were contacted for more information on the development of the tool or for access to the relevant portion of the manual, unsuccessfully. A search was done to find other studies that reported the development of the measure or studies in which the CARS was used. One such study was identified (Schopler et al., 1980) which reported an internal consistency coefficient of $\alpha = .94$ and interrater reliability of .71. Two other studies (DiLalla & Rogers, 1994; Breidbord & Croudace, 2013) were also identified: DiLalla & Rogers (1994) presented the results of an exploratory factor analysis of the CARS while Breidbord & Croudace (2013) examined the interrater reliability and internal consistency from various studies. Based on the results of these studies (Schopler et al., 1980; DiLalla & Rogers, 1994; Breidbord & Croudace, 2013) and evidence from McConachie et al. (2015), internal consistency, structural validity and reliability were rated as moderate. The overall rating for the measure was medium based on COSMIN guidelines.

Additionally, as per the publisher's guidance, some training and specific educational qualification are required before using the CARS. Thus, it seems inappropriate for further consideration for screening adolescents in Africa.

Developmental Behavior Checklist-Autism Screening Algorithm (DBC-ASA). The DBC-ASA (Brereton et al., 2002) is a 29-item autism screening measure derived from the Developmental Behavior Checklist (DBC). The DBC was revised and updated to the DBC2 in 2018. The parent version of the DBC is available in the following languages: Chinese, Arabic, Croatian, Dutch, French, Finnish, German, Greek, Hindi, Norwegian, Portuguese (Brazilian), Italian, Japanese, Spanish, Swedish, Turkish and Vietnamese.

Deb et al. (2009), screened a total of 109 children aged 3 – 17 years with intellectual disabilities for autism using the instrument. Forty-four of the children were between 3 – 9 years old, 50 of

them between 10 – 15 years old and 15 participants were older than 15 years. A cut-off score of 19 for the 3 – 9 years olds yielded a sensitivity of 1 and specificity of .71 while a cut-off of 26 for the 10 – 15-year-olds yielded a sensitivity of .70 and specificity of .75. When a total population cut-off score of 20 was applied, sensitivity was .90 and specificity .60. The figures generated by Deb et al. (2009) differ from those obtained in Brereton et al. (2002) where a cut-off score of 14 yielded sensitivity of .86 and specificity of .55 and a cut-off score of 17 yielded a sensitivity of .79 and specificity of .63. Perhaps this could be attributed to the characteristics of the participants as noted by Deb et al. (2009); they screened for autism in children with intellectual disabilities only, while the Brereton et al. (2002) examined the validity of the tool among individuals with and without intellectual disabilities. Neither study reported a PPV or NPV. There was no validation against an accepted gold standard tool; rather, the participants received a clinical diagnosis of autism based on the ICD-10-DCR (International Classification of Diseases 10th Revision, Research Diagnostic Criteria) in the Deb et al. (2009) study.

Appraising the quality of the reviewed study (Deb et al., 2009), the content validity, structural validity, cross-cultural validity, internal consistency, construct validity and reliability were all rated as undetermined. Criterion validity was rated as sufficient based on the evidence. As peer-reviewed studies do not always provide sufficient information, the authors of the DBC were contacted to confirm which of the validities were examined. Based on the evidence provided by the authors and excerpts from the manual, reliability, internal consistency, convergent validity, criterion validity, discriminative validity, and concurrent validity were all rated as positive. Since the DBC-ASA is not an independent measure but an algorithm within the DBC, the relevant psychometric (discriminative validity) property of the DBC-ASA was assessed. Brereton et al. (2002) and Deb et al. (2009) both reported that the DBC-ASA had very good discriminative ability. However, there remains inadequate information on cross-cultural validity, placing a limitation on its use in an African context. The overall rating for the tool based on the COSMIN checklist was medium.

Pervasive Developmental Disorder in Mentally Retarded Persons (PDD-MRS). The PDD-MRS is a 12-item questionnaire designed for clinician screening for autism amongst those with intellectual disabilities. It has dichotomous items spread across the following domains: communication, social behaviour and stereotyped behaviour. It was designed to be used with

children and adults ages 2 – 55 years old. The original Dutch version: the Autisme- en Verwante kontaktstoornissenschaal voor Zwakzinnigen (AVZ) was developed specifically for use with people with intellectual disabilities (Kraijer, 1990) with a revision in 1994 (Kraijer, 1994). The instrument is based upon the DSM-III-R criteria for pervasive developmental disorders and has been widely used in the Netherlands and Belgium.

Kraijer & de Bildt (2005) described and discussed the construction of the scale and its validation. The psychometric properties were tested on a sample of 1,230 participants with varying levels of intellectual disabilities. The resulting sensitivity at a cut-off score range of 10 – 19 was .92, while specificity was .92, but neither the PPV nor NPV was reported. Internal consistency for participants with functional speech was reported as α = .86 and for those without speech α = .81. Cortés et al. (2018) developed and validated the Escala de Valoración del Trastorno del Espectro Autista en Discapacidad Intelectual (EVTEA-DI), the Spanish version of the PDD-MRS. Reported results were r = .78 for convergent validity between the EVTEA-DI and the CARS, internal consistency measured by the Kuder-Richardson-20 (KR-20) was .71. At a cut-off score of 30, sensitivity was .71, specificity of .90, PPV = .73 and NPV = .90. To assess the discriminative validity of the EVTEA-DI, Cortés et al. (2018) utilised the Youden Index (YI). At a cut-off score of 8, sensitivity = .84 and specificity = .83.

For the PDD-MRS, content validity was rated as moderate based on the evidence from reviewed studies. Structural validity, internal consistency, criterion validity, and construct validity were all rated as positive as there was sufficient methodological evidence found to support the rating. There was moderate evidence for cross-cultural validity since individuals with varying disabilities from different populations were participants. Studies were completed with Dutch and Spanish speaking participants. Reliability was rated as insufficient based on the COSMIN rating of lowest score counts. Authors were contacted for further evidence without success. The overall COSMIN rating for this tool was medium.

Child Behavior Checklist (CBCL). The CBCL (Achenbach & Rescorla, 2001) is now a component of the Achenbach System of Empirically Based Assessment (ASEBA). The CBCL is a caregiver report questionnaire on which children and teenagers (2-18 yrs) are rated for various behavioural and emotional difficulties. Associated with disorders from the DSM-5, it measures difficulties on a scale made up of eight categories – rule-breaking behaviour,

anxious/depressed, social problems, somatic complaints, thought problems, attention problems, withdrawn/depressed and aggressive behaviour. The form consists of 118 items that take between 30 minutes to an hour to complete. The CBCL has been translated into 60 different languages. Previous versions of the checklist were not designed to screen for autism in young children older than 4 years, and 6 years in the current revision (Mazefsky et al., 2011).

However, Ooi et al. (2011) aimed to derive and test an autism scale that could significantly differentiate children and adolescents with and without autism using the CBCL. The study participants were between 4 and 18 years old. The researchers considered whether eight scale factors could significantly differentiate individuals with and without autism, and they reported a sensitivity range of 48 – 78% and a specificity range of 59 – 87%. Following this, Ooi et al. (2011) derived and tested an autism scale comprised of items taken from the CBCL that significantly differentiated autistic children from other groups. Results showed that nine specific items were predictive of autism with sensitivity ranging from .68 – .78 and specificity range of .73 – .92. The PPV and NPV were not reported . The CBCL scores falling below the 93rd percentile are considered normal, scores between the 93rd to 97th percentile are borderline clinical, while scores above the 97th percentile are in the clinical range. Results of Ooi et al. (2011) are consistent with findings from previous studies (Mazefsky et al., 2011). Both Ooi et al. (2011) and Mazefsky et al. (2011) reported that the CBCL scales with more effective discriminative abilities between the typical and autistic school-aged children were the 'Thought Problems, Social Problems and Withdrawn/Depressed' categories.

Regarding the quality appraisal from the reviewed study Ooi et al. (2011), the content validity for the CBCL was rated as indeterminate while structural validity was rated as positive, given the quality of the evidence reviewed. Criterion validity, construct validity and internal consistency were all rated as undetermined as there was not sufficient evidence. There was moderate evidence for reliability, with sufficient evidence to rate cross-cultural validity as positive. The scale which was originally developed in English was used with participants in three different languages (English, Malay and Tamil) and five different groups (Ooi et al., 2011). The authors were contacted for more evidence or access to relevant portions of the manual. Based on the author's response, content validity, reliability, criterion validity, construct validity, internal consistency and discriminative validity were all rated as sufficient. The overall rating for the

CBCL was, medium, based on the level of evidence using the COSMIN checklist. Although work has gone into translating the tool into different languages and deriving a potential autism specific screening subscale from the CBCL, some training is required. The level of training depends on how the data are to be used. For LMICs such as Nigeria, Ghana, Kenya and other African countires, these requirements are potential barriers.

Mobile Autism Risk Assessment (MARA). Duda et al. (2016) described the MARA, a new 7-item parent or caregiver questionnaire designed to screen for individuals at risk of autism. The MARA was developed based on the analysis of a pool of ADI-R score sheets of individuals with and without autism. An alternating decision tree algorithm was used to generate the questions and responses. The tool is administered and scored electronically, and the reported sensitivity was .90 and specificity was .80. Given that the data used for testing the measure were taken from the ADI-R, it should follow that the discriminatory ability and construct validity should be good. The MARA was validated against the Autism Diagnostic Observation Schedule (ADOS) and the PPV was .67, and NPV was .95. Duda et al. (2016) reported no specific cut-off scores; however, they referenced Wall, Dally, Luyster, Jung, & DeLuca (2012) where they used a categorical variable with two options – autistic or not autistic. Although the MARA looks promising, more large-scale reliability and validity studies with participants of differing developmental abilities are needed.

Based on the reviewed study (Duda et al., 2016), there was adequate evidence to rate structural validity as positive. Internal consistency, reliability, criterion validity, and construct validity were rated negative due to insufficient evidence. Content validity was rated as insufficient as the involvement of experts and users was unclear. Evidence for cross-cultural validity was insufficient and was rated as very low. The authors were contacted for more information and possible access to the manual if available. Based on feedback from one of the authors, content validity was revised to a positive rating. However, other studies provided were not on the MARA but on detecting ASD through Machine Learning. Participants in those studies were children younger than 5-years of age, thus not meeting the inclusion criteria for this review. Based on the COSMIN standard, the overall rating for the measure was low. Also, using this tool in Africa could be challenging, given that not everyone has internet access or personal computers.

EDUTEA: A DSM-5 teacher screening questionnaire for autism & social communication disorders (EDUTEA). The EDUTEA was developed in Spain as a brief autism screening tool for use by teachers and school professionals who had limited time (Morales-Hidalgo et al., 2017). The EDUTEA is an 11-item questionnaire based upon DSM-5 diagnostic criteria and was designed to enable teachers to gain information about the social interactions, behaviours and communication skills of children. The tool was validated against the ADOS-2, ADI-R and compared to the CBCL, Childhood Autism Spectrum Test (CAST) and Schedule for Affective Disorders and Schizophrenia (K-SADS-PL). Scoring of items is on a 4-point Likert scale, resulting in a minimum score of 0 to a maximum score of 33.

In evaluating the discriminatory ability and psychometric properties of the tool, Morales-Hidalgo et al. (2017) recommended a cut-off score of 10. At the recommended cut-off, the EDUTEA successfully discriminated between those with autism and related disorders and those with ADHD with an associated sensitivity of .83 and specificity of .73. For differentiating individuals at risk of autism or social pragmatic communication disorder (SCD), the authors reported good discriminatory abilities at the cut-off score of 10, with sensitivity = .87 and specificity = .91NPV of .99 and a PPV of .87. The two-factor internal consistency for the measure was α = .95 for social communication impairments and α = .93 for restricted behaviour patterns. Overall internal consistency was α = .97. No other studies using the instrument were found from the literature search.

Content validity was rated as positive as teachers were involved in the development of the instrument. The structural validity, internal consistency, criterion validity, reliability and construct validity were all positive with moderate evidence. However, cross-cultural validity was judged as having insufficient evidence. The overall rating based on COSMIN standards was medium.

Autism Diagnostic Inventory-Telephone Screening in Spanish (ADI-TSS). Vranic et al. (2002) developed the ADI-TSS as a semi-structured interview administered over the telephone. ADI-TSS was modelled upon the Autism Diagnostic Interview-Revised (ADI-R) with forty-seven questions in three areas. The final version used in the study was administered to 59 participants and had a sensitivity of 1, and a specificity of .66 with no PPV or NPV reported. Although this tool was developed over fifteen years ago, no other studies validating its use and

properties were found. Interrater reliability for the subscales were as follows: social reciprocity α = .94, verbal communication α = .93, non-verbal communication α = .94, and repetitive behaviour α = .94.

Content validity for the subscales was rated positive, while the overall content validity was rated low due to insufficient evidence for end-user input in the development of the tool. Structural validity and internal consistency were rated insufficient. Cross-cultural validity was rated insufficient as the translation methodology was unclear. Although inter-rater reliability for the subscales was shown, there was insufficient evidence for the reliability of the total tool; thus, this was rated insufficient. Based on the COSMIN checklist, the tool was rated as low overall. The feasibility of using the ADI-TSS in Africa, where there are high costs associated with mobile telephone use would be a challenge.

Diagnostic Behavioral Assessment for Autism Spectrum Disorder-Revised (DiBAS-R). The DiBAS was developed by Sappok and colleagues (2014b) to help with screening autism amongst adults with intellectual disabilities. It was designed to be administered by caregivers or individuals knowledgeable about the person, but who also lacked specific knowledge about autism. The 20-item questionnaire was derived from the ICD-10 and DSM-5 criteria for autism. To improve its diagnostic validity further, a single item was deleted following a pilot study and item-revision of the DiBAS (Sappok et al., 2014a). The resulting 19-item screening tool can be completed in 5-minutes by a caregiver, family member, staff or any person who is familiar with the individual.

Heinrich et al. (2018) assessed the diagnostic validity of the DiBAS-R in 381 adolescents and adults with intellectual disabilities, some of who had autism. Study participants ages ranged between 16 – 75 years. Based on the recommended cut-off score of 29, the reported results were sensitivity = .82, specificity = .67, the PPV = .44 and the NPV = .92. The participant's diagnosis was confirmed using the ADOS and ADI-R.

Based on the reviewed study (Heinrich et al. (2018), content validity was rated as undetermined. Expert clinicians participated in the development, but the item reduction process was unclear. Assessment of comprehensibility and comprehensiveness was also unclear. Evidence for cross-cultural validity, structural validity and internal consistency were also insufficient. Reliability was rated as insufficient, while criterion validity and construct validity had sufficient evidence to

rate them positive. The authors were contacted for access to the manual or further evidence on the tool's development. Since the manual is in German, the authors provided Sappok et al. (2014a) in which the relevant information was reported. Following this, the content validity, structural validity, internal consistency, reliability, convergent and discriminative validity were all rated positive. However, evidence for cross-cultural validity remained insufficient. DiBAS-R was rated as medium based on the additional evidence using the COSMIN checklist. DiBAS-R is currently available only in German, thereby limiting the feasibility of using it in Africa.

Adapted Autism Behaviour Checklist (AABC). The AABC which is based on the Autism Behavior Checklist (Krug et al., 1980) is a 57-item measure developed in Turkey by Özdemir & Diken (2018). Modifications were made to the original form to include the ICD-10 and DSM-5 criteria for autism. The measure was designed to be completed by a parent, primary caregiver, or a teacher familiar with the individual and then scored and interpreted by a trained professional.

Özdemir & Diken (2018) assessed the diagnostic validity of the AABC in 1,133 children and adolescents with autism and intellectual disabilities. Study participants ages ranged between 3 – 15 years. Reported results were r = .73 between the AABC and the Gilliam Autism Rating Scale-2 Turkish Version (GARS-2 TV), internal consistency measured by the Kuder-Richardson-21 (KR-21) was .89, test-retest reliability was r = .82 and correlation between the two-factors (social limitations and problematic/repetitive behaviours) was r = .46. At a cut-off score of 13, the measure discriminated between the ASD and ID groups reliably with a sensitivity of .87 and specificity of .82.

Based on the COSMIN checklist, content validity, structural validity, internal consistency, reliability, criterion validity and construct validity were all rated as positive. Cross-cultural validity was rated as insufficient based on the evidence. The tool has only been used in Turkey. Since this measure is only available in Turkish, the feasibility of using it in Africa is limited as substantial resources would be required for translation. The overall rating for the measure was medium.

Intellectual Disability

The ninteen studies identified focused upon people with intellectual disabilities and included a total of 3,129 participants with age ranging from 3 to 74 years. Like autism, studies with

participants younger than 11-years or older than 26-years old were included when some or the majority of their participants were within the specified age range of the inclusion criteria. The number of studies by country was as follows: UK (n = 7), USA (n = 4), Norway (n = 5), and one each from Australia, Netherlands, and Belgium (Table 9). Three of the studies (McKenzie et al., 2012b; Trivedi, 1977; Ford et al., 2008) involved adolescents only while fifteen studies involved a combination of children, adolescents, and adults. The screening tools used in the studies were the Slosson Intelligence Test (SIT), Learning Disability Screening Questionnaire (LDSQ), Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q), Screener for Intelligence and Learning Disabilities (SCIL), Hayes Ability Screening Index (HASI), and the Quick Test (QT). Validation of these screening tools was against full-length tests considered as the gold standard, such as the), different editions or versions of the Weschler scales and the Adaptive Behavior Assessment System-Second Edition (ABAS-II). In some instances, screening tools were compared to other full length scales, such as the Stanford-Binet Intelligence Scale, or to similar short measures. For example, the HASI was compared with the KBIT and the SIT with the Peabody Picture Vocabulary Test (PPVT).

The quality ratings for the included studies are found in Tables 6 (in Appendix) and 7. Five of the included studies made use of a single group of participants, while eight used a between-group design, and six a within-group design. Each screening tool is considered in turn below.

Intellectual Disabilities Screening Tools

Slosson Intelligence Test (SIT) and Slosson Intelligence Test-Revised (SIT-R). The original SIT was developed by Richard Slosson in 1963 and used as part of an assessment to determine whether an individual has an intellectual disability, measured as IQ. At the time of this review, no studies utilising the third and fourth versions of the SIT were found. Rotatori and Epstein (1978) assessed the ability of special education teachers without previous psychological testing experience, to reliably administer the SIT. Reported test-retest reliability results (r = .94) appeared excellent, indicating that the test was reliable over time when administered by special education teachers. To examine the concurrent validity of the revised SIT, Kunen et al. (1996) compared the SIT R to the Stanford-Binet Intelligence Scale, Fourth Edition. The correlation was high (r = .92), but the consistency of the IQ classification between the two instruments for those who had intellectual disabilities was poor. In comparison to the Stanford-Binet Intelligence

Scale, the SIT-R had insufficient evidence of construct validity due to discrepancies in match rates between the SIT and the Stanford-Binet. For instance, for the entire study sample with IQs ranging from 36 to 110 (Kunen et al., 1996), there was a 50% match rate between the Stanford-Binet and the SIT for all the classifications, mild, moderate, average and low average IQs. Nevertheless, for the mild to moderate categories out of the 38 participants categorised as mild on the Stanford-Binet, SIT categorised them as 1- low, 2- slow, 9- mild, and 26- moderate. Trivedi (1977) meanwhile examined the comparability of the SIT against the Wechsler Intelligence Scale for Children (WISC) in adolescents. He found significant correlations between the WISC and SIT when compared on mental age (r = .87) and IQ (r = .86). Trivedi (1977) concluded that the SIT reliably approximated the WISC as a screening tool.

Blackwell and Madere (2005) commented that the SIT-R demonstrated and fulfilled its stated purpose of "being a valid, reliable, individual screening test of general verbal cognitive ability" (p. 184) but have also suggested problems with the reliability and validity of the SIT-R. Reviews by other authors have also raised concerns about the reliability and validity of the SIT-R (Campbell & Ashmore, 1995). Potential challenges regarding the use of the SIT-R with those from multicultural backgrounds, or where English is a second language were reported by Blackwell & Madere (2005). Other limitations of the SIT-R are its inability to measure functioning levels of other intellectual areas such as perceptual-motor functioning. There is also the difficulty of comparing SIT scores with those of other IQ tests for persons older than 16-years of age due to the unclear and insufficient methodological information given by the developers (Campbell & Ashmore, 1995). Although the SIT has the above limitations, one advantage is that persons with limited psychometric training and knowledge can administer it.

Based on the COSMIN checklist, the SIT (or SIT-R) was rated as low overall. There was sufficient evidence for reliability from the studies reviewed for it to be rated as moderate. Content validity and structural validity were rated undetermined. Both criterion validity and construct validity were rated as inconsistent. Internal consistency and cross-cultural validity were rated as negative, based on the poor amount of evidence.

Quick Test (QT). The QT is an intelligence test measuring verbal information processing and receptive vocabulary (Ammons and Ammons, 1962). It comprises three parallel forms with 50 items, each of which can be administered to children and adults. Verbal intelligence is measured

by the ability to match words of increasing complexity to pictures. Sawyer & Whitten (1972) investigated the concurrent validity of the individual and combined scores of QT against the WISC sub-tests. Moderate correlations (r = .33 – .52) were reported for the picture arrangement, coding, performance scale score and the full-scale score. For the verbal scale, the correlation between both the QT and the WISC was between r = .31 and .34 for both the individual and combined forms of the QT. One challenge with the QT is that it predominantly measures verbal skills. This limitation may have impacted the Sawyer & Whitten (1972) study, as most of the participants had limited verbal ability. Moreover, the pictures used are rather old-fashioned and may not transfer well to the African context.

Based on the COSMIN checklist, the overall evidence for the QT was very low. Structural validity, internal consistency and reliability were rated low based on insufficient amount of evidence both from the study and manual. There was sufficient evidence to rate the construct validity, content validity and criterion validity as positive, while cross-cultural validity was undetermined. The overall rating for the QT was very low.

Hayes Ability Screening Index (HASI). The HASI is a brief screening tool for intellectual abilities comprised of four subtests covering background information, puzzle, clock drawing and backward spelling (Hayes, 2000). The HASI has been used predominantly in criminal justice settings to identify vulnerable persons with intellectual disabilities. HASI is designed for use with people aged 13 to adulthood. For those aged 13 – 18 years, the cut-off score is 90, while for those older than 18-years, it is 85. Some training is required before its use.

Hayes (2002) reported on the construct validity of the HASI and the correlation with the Kaufman Brief Intelligence Test (KBIT), Wechsler Abbreviated Scale of Intelligence (WASI) and WISC-III. The total population sample correlation between the HASI and KBIT was reported as high (r = .62). The reported sensitivity for the study was .82, and the specificity was .72. Hayes (2002) suggested that the youth cut-off be maintained at 90. A different study (Ford et al., 2008) which had all adolescent (10 – 19-year-olds) participants, found a correlation of r = .55 between the HASI and the FSIQ of the WISC-IV or the Wechsler Adult Intelligence Scale (WAIS-III) and r = .38 with the Vineland Adaptive Behavior Scale (VABS). At the recommended cut-off score of 90 for those below 18 years of age and 85 for those over 18 years old, the authors reported a poor agreement (k = .25) between the HASI and the FSIQ from the

Weschler scales when categorising as ID. Sensitivity at these cut-off scores was .66 and specificity of .51. Lowering the cut-off score to 80.2 yielded better agreement (k = .54) a sensitivity of .80 and specificity of .65. Søndenaa et al. (2007) translated the HASI to Norwegian and validated the construct and criteria of the screening tool against the Norwegian version of the WAIS-III. The study participants were between 17 and 60 years old. The authors found a high correlation between both instruments (r = .81) with an internal consistency of α = .76. Søndenaa et al. (2007) also reported that scores on all HASI subtests, WAIS-III FSIQ and the verbal and performance subscales were significantly correlated with r above .61. At the recommended cut-off score of 85 for indicating ID, the sensitivity was 1 and specificity .57. However, Søndenaa et al. (2007) adjusted the cut-off score to 81 for their sample to reduce the over-inclusion of false positives. The alternative cut-off of 81 yielded a sensitivity of .95 and specificity of .72.

In the Søndenaa et al. (2008) prevalence study, the HASI was validated against the WASI as a screening tool. The HASI was found to be somewhat overly inclusive with a specificity of 72.4% and sensitivity of 93.3%. Correlations between the WASI full-scale and HASI were significant with r = .72, verbal tests r = .63 and performance tests r = .74. In Søndenaa et al. (2011), the criterion validity of the HASI was examined against the WASI with a psychiatric population. The study reported the over categorisation by the HASI with a sensitivity of 1 and specificity of .35 at the recommended cut-off score as previously mentioned. However, the authors argued that the HASI is designed to be overly inclusive, since it is better to identify everyone who may need full assessments, rather than miss some people. Also, Søndenaa et al. (2011) reported moderate correlations between the subtests of the WASI and HASI (r = .67). However, when the "background information" subtest was eliminated, correlation increased to r = .71 and internal consistency of α = .67.

To et al. (2015) examined the discriminative and convergent validities of the Dutch version of the HASI against the WASI-III in persons with substance abuse problems. Convergent validity between the HASI and WAIS-III FSIQ scores, were significantly correlated (r = .69). There was also a correlation between the HASI subtests and the WAIS-III as follows: background information r = .58, spelling r = .50, puzzle r = .46, clock drawing r - .45, verbal subscale r = .70, and the performance subscale was r = .63. Discriminant validity was reported as significant from the receiver operating characteristics (ROC), with an area under the curve (AUC) of .95

yielding a sensitivity of .91 and specificity of .80 at the cut-off score of 85. In Braatveit et al. (2018), the convergent and discriminative validities of the Norwegian version of the HASI were examined in a population of persons with a substance abuse history. At the cut-off of 85, sensitivity was reported as 1 and specificity of .65. Braatveit et al. (2018) also reported that lowering the cut-off score to 80.7 yielded increased specificity of .81 without affecting the sensitivity. Similar to Søndenaa et al. (2011), Braatveit et al. (2018) also mentioned that the over-categorisation by the HASI was intended to be a means of detecting other persons with/without intellectual disabilities but who may benefit from further evaluation. Regarding convergent validity, Braatveit et al. (2018) correlated the HASI against the full-scale WAIS-IV with a significant correlation ($r = .70$).

Based on the reviewed studies, and the COSMIN checklist, the overall rating for the HASI was low. Reliability was rated as negative due to insufficient evidence. Structural validity had inadequate evidence and was rated as undetermined. Content validity was rated as low due to insufficient evidence. Criterion validity and construct validity were rated positive with excellent evidence. There was moderate evidence for a positive rating on the cross-cultural validity based on the use of the Norwegian and Dutch versions, as well as the original Australian version. To ensure that all relevant properties of the tool were properly rated, the manual was consulted. Based on the manual, additional ratings employing the COSMIN were made. Content validity remained low as there was no evidence on expert clinicians or end users involvement in the development. Evidence for internal consistency was not in the manual thus a rating of insufficient was given. Reliability was rated as insufficient, while criterion validity and construct validity had sufficient evidence to retain their positive rating. The overall rating of the HASI was revised to medium following the combined evidence from the studies and the manual. Although the HASI has been adapted for use in two further languages and environments outside of the original development area, most of the studies used the tool in the Criminal Justice System. Studies that employed the tool with adolescents outside of the CJS would have been more useful for forming a decision on adapting it for use in Africa and countries like Nigeria.

Learning Disability Screening Questionnaire (LDSQ). McKenzie and Paxton (2006) developed this 7-item screener for the identification of adults with intellectual disabilities to assist in deciding eligibility for community services. The LDSQ has also been used in criminal

justice and forensic settings. Areas assessed include literacy, living situation and employment. The LDSQ has been reported to have both criterion and convergent validity when compared to the WAIS-III (McKenzie and Paxton, 2006). McKenzie et al. (2012a) examined the convergent and discriminative validities of the LDSQ in forensic settings. Convergent validity between the FSIQ and the LDSQ was reported as highly significant with a correlation coefficient of $r = .71$. The authors also reported good discriminative ability of the LDSQ with a sensitivity of 82.3% and specificity of 87.5% based on the receiver operating characteristics analysis (AUC = .898). PPV and NPV were reported as 92.9% and 73.7% respectively. McKenzie et al. (2015) validated the LDSQ's criteria against a standardised tool, the WAIS-IV FSIQ and reported a good correlation between them with a sensitivity of .92 and specificity of .92 (AUC = .945). Convergent validity was reportedly significant for the WAIS-IV FSIQ and LDSQ total performance with a coefficient of $r = .71$. Significant correlations were also reported for the subtests – verbal comprehension ($r = .54$), perceptual reasoning ($r = .69$), working memory ($r = .58$), and processing speed ($r = .58$). Although these studies by McKenzie et al. (2012a; 2015) reported excellent psychometric properties for the LDSQ, the independent study by Stirk et al. (2018) reported a sensitivity of .67 and specificity of .71 at the threshold given by McKenzie et al. (2015), showing that the LDSQ may require more investigation to align the properties with recommended standards (Glascoe, 2005).

Based on the evidence from the studies reviewed, criterion validity and construct validity were rated as moderate. Content validity was rated as insufficient since there was not enough evidence of user participation in the development of the tool. Structural validity, internal consistency, cross-cultural validity, reliability, were all rated low due to insufficient evidence. The manual was obtained to confirm which of the tool's properties were examined during development. From the manual, there was moderate evidence for content validity, discriminative validity, and convergent validity. Interrater reliability was assessed while there was no evidence for internal consistency. Combining the evidence from the studies and the manual, the overall quality of the LDSQ was rated as medium using the COSMIN checklist. Like the HASI, this measure has been used primarily with adults in the CJS and forensic services. However, unlike the HASI, evidence to support the cross-cultural application was not apparent, and so the feasibility of its use with African adolescents is limited.

The Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q). The CAIDS-Q, modelled after the LDSQ, was developed by McKenzie and Paxton in 2012 as a short 7-item screening questionnaire for detecting intellectual disabilities in children and adolescents in mental health and forensic services. It is designed for use with individuals aged 8 – 18 years. According to McKenzie & Paxton, the CAIDS-Q can discriminate between those with and without intellectual disabilities with 97% accuracy. Four studies, McKenzie et al. (2014; 2012b, 2012c; 2019) evaluated and validated the psychometric properties of the tool and reported values within recommended standards. McKenzie et al. (2014) assessed the discriminatory ability of the CAIDS-Q against a short form of the WISC-IV with a sample of children aged 10 to 11 years with and without intellectual disabilities (who had been fully assessed for this on either the WISC IV or the WAIS III). Overall, the WISC-IV short form itself led to the correct classification of 91% of the participants. When broken down, the classification of those with intellectual disabilities was 92% correct while those without intellectual disabilities was 91% correct, using the WISC-IV short form. AUC for the WISC-IV was .98 which gave a PPV of .87 and NPV of .95. The CAIDS-Q led to the correct classification of 89% of children with intellectual disabilities, and 88% of those without; the PPV was .92, and the NPV was .85 based on an AUC = .94. Overall, the CAIDS-Q correctly classified 88% of the participants. McKenzie et al. (2012c) evaluated the face, construct, criterion, convergent and discriminative validity of the CAIDS-Q with comparisons made to either the WISC-IV FSIQ or the WAIS-III FSIQ depending on the participant's age. Results obtained from the study showed high internal consistency ($\alpha = .88$), significant correlations between the CAIDS-Q and the WISC-IV FSIQ ($r = .78$), and significant correlations between the CAIDS-Q and the WAIS-III FSIQ ($r = .79$). At a cut-off of 62 for the children (8 – 11 years), the measure had a sensitivity of .97 and specificity of .86, and at a cut-off of 64 for the adolescents (12 – 18 years) the sensitivity was .96 and specificity .85. McKenzie et al. (2012c) reported that there was no significant difference between age and the CAIDS-Q score for the total population ($r = .02$). McKenzie et al. (2012b) evaluated the convergent and discriminative validity of the CAIDS-Q against the WISC-IV in a forensic setting. Reported outcomes were significant correlations between the CAIDS-Q and the FSIQ ($r = .76$), with correlations between the CAIDS-Q score and the subtests as follows: verbal comprehension ($r = .54$), perceptual reasoning ($r = .65$), working memory ($r = .52$), and processing speed ($r = .74$). Other results include a PPV of 1, NPV of 1 and good internal

consistency (α = .72). McKenzie et al. (2019) examined the convergent validity, test-retest reliability, interrater reliability, sensitivity, specificity, PPV and NPV in a paediatric neurodevelopmental setting based on previously determined cut-off scores. Convergent validity of the CAIDS-Q was examined against the WISC-IV and Adaptive Behaviour Assessment System, Second/Third Edition (ABAS II/III). Reported correlations between the CAIDS-Q and the FSIQ ranged between r = .62 - .79, with correlations between the CAIDS-Q and the ABAS GAC ranging between r = .48 - .60. Other results include a PPV of 1, NPV of .78, sensitivity = 1 and specificity = .88 for the total sample. A two-week time frame yielded a test-retest correlation of r_{32} = .90 while interrater reliability k was between .26 and 1 for the four items (time, read, write and laces) tested.

From the studies reviewed, content validity was undetermined as user participation in the development was unclear. Assessment of comprehensibility was also unclear. Structural validity was rated as negative due to insufficient evidence. Evidence for reliability was moderate from the studies. Cross-cultural validity was rated as moderate since the measure was used with two different age groups: children and adolescents. Criterion validity was also rated as moderate. However, incorporating information from the manual, content validity, structural validity, internal consistency, reliability, criterion validity and construct validity were all rated as positive with moderate evidence. Based on the COSMIN checklist, the overall rating of the tool was medium.

Screener for Intelligence and Learning Disabilities (SCIL). The SCIL is a tool for identifying persons with a level of general intellectual functioning that falls within and below the "borderline" range (Nijman et al., 2016). The SCIL comprises elements of social adaptive skills, language comprehension, education, arithmetic, reading and writing abilities. Geijsen et al. (2016) examined the predictive validity of the SCIL for identifying intellectual disabilities amongst adolescents and adults in a criminal justice setting (police detention). Reported results from the study showed that the SCIL total score correlated moderately with the WAIS-III short form (r = .56) with a sensitivity of .72 and specificity of .70. The PPV and NPV were not reported. Additionally, reliability was reported as α = .64 and α = .84 in a previous study (Kaal, Nijman, & Moonen, 2015). Nijman et al. (2016) conducted two further studies, split into adults and adolescents, and investigated the predictive validity of the SCIL. Participants in both groups

included persons with and without intellectual disabilities. At the suggested cut-off of 19, sensitivity was .82 and specificity was .89 for adults, and for adolescents the suggested cut-offs varied according to age. For those aged between 16 and 17-years old, a cut-off score of 18 resulted in a sensitivity that was .80 and specificity that was .84; for those aged 14 to 15-years old, a cut-off score of 16 resulted in a sensitivity of .85 and specificity of .82. The AUC for adolescents as a total group was .91, and .93 for adults. The SCIL had test-retest reliability of r = .92. Nijman et al. (2016) analysed the internal consistency using the split-half method which yielded high correlations; α = .84 in the first half and α = .82 in the second half.

Based on the COSMIN checklist, the overall rating for SCIL was moderate. Content validity was rated positive with moderate evidence based on the involvement of experts and end-user. Structural validity, criterion validity and internal consistency were all rated as highly positive with enough evidence. Reliability had moderate evidence with consistent findings in both studies. Cross-cultural validity was positive as participants were recruited from different cities, police stations and refugee sites. The SCIL showed promising results, but more studies to validate the tool are required.

Discussion

Identification and selection of a user friendly, accessible, time-efficient, and useful screening tool for use in Africa requires careful thought and consideration. The focus of this review was on identifying potentially useful tools for screening African adolescents and younger adults with autism and/or intellectual disabilities. This age range was the focus as many who have intellectual disabilities and/or autism are noticed as they become more independent and when they begin to interact more often with others outside of their immediate family, for example in secondary school settings and the wider community. In interacting with these environments, disabilities and challenges become more obvious. Such adolescents may not have received a diagnosis earlier in their life because of a lack of awareness, insufficient or inadequate resources, limited numbers of professionals and the families sometimes not seeking immediate help for those individuals with autism or intellectual disabilities till later in life. To begin to address this gap, appropriate and suitable screening tools need to be designed or identified for use in Africa. Towards this, the aims of this review were to (1) describe and critically appraise a range of short screening tools for the detection of intellectual disabilities and autism, (2) consider the

psychometric properties of these tools, and (3) consider the appropriateness of using these tools across a range of cultures. A discussion of the review findings is presented below.

Description and appraisal of short screening tools

A total of 12 tools screening tools for autism were identified through this review. The tools are the ADI-TSS, EDUTEA, PDD-MRS, DiBAS-R, AQ-10, ASSQ-REV, SCQ, CARS, CBCL, DBC-ASA, AABC and the MARA. Apart from the AQ-10 adolescent version, all the other tools were designed to be used across a wide age range. The CARS and CBCL were not originally developed as screening tools, but the studies (Mesibov et al., 1989; Ooi et al., 2011) reviewed utilised them as such with the intent of developing subscales for autism screening. Moreover, the CBCL has over 100 items and takes between 30 minutes to an hour to complete, which goes against the timing for brief tools. Both the CARS and CBCL require some training and specific qualifications before use. Given the socioeconomic climate of African countries, the resources required to gain specific administrative qualifications for these tools may not be readily and widely available. As such, there will be challenges associated with routine use within Africa.

Both the PDD-MRS and DiBAS-R can be used across a wide age range from 2 – 80-year-olds and administered in 5 to 20 minutes. The wide age range allows for their use with adolescents while the short administration time qualifies them as short and time-efficient tools. The PDD-MRS and DiBAS-R were designed for use with persons who are known to have intellectual disabilities. Limiting the measures to those with known intellectual disabilities presumes those individuals have been diagnosed; this is not entirely the situation in Africa. Considering this design limitation, the feasibility of their use in Africa will be challenging.

The MARA, ADI-TSS and SCQ were modelled after the ADI-R. While the SCQ and MARA take between 5 – 10 minutes to administer, the ADI-TSS takes between 20 – 40 minutes. The lengthy administration time of the ADI-TSS may be because of the telephone administration. As an over-the-telephone screening tool, the usefulness of ADI-TSS in Africa, where not everyone may have access to a telephone, poses immediate limitations. Similar constraints are associated with the MARA, which is a computer-based parent or carer administered screening tool. The number of persons with immediate access to either a smart device, personal computer or constant electricity is likely to be low in the African continent or individual African countries. This lack of immediate access to smart devices poses a limitation to the usefulness of the MARA in Africa.

Meanwhile, for the SCQ, seven (32%) out of the 22 studies reviewed employed this tool and this observation is consistent with findings from previous studies that concluded the SCQ was used more widely in research (Bozalek, 2013). The SCQ comprises two forms: lifetime and current. The SCQ current form is used to assess an individual's behaviour during the past 3 months while the lifetime form assesses the developmental history. One advantage of the SCQ is the availability of the lifetime form, which enables information gathering for adolescents who have never been screened. This feature, amongst others, makes the SCQ a viable option for use with African adolescents.

The DBC-ASA, which is a subset of the Developmental Behavior Checklist, is limited to those under age 18-years, which is about the midpoint of the adolescent age range (11 – 26 years). The upper age limit of the screening tool poses a current challenge for routine use of the tool. On this basis, adopting the tool for use in Africa does not seem practicable without further standardisation work inclusive of a wider age range.

Both the ASSQ and AQ were developed for persons with HFA. While one of the reasons tools are developed is to bridge a gap or meet a need, in the African setting where screening is still in its infancy, using such disability-specific tools will not yield optimal results. EDUTEA, an 11-item questionnaire, was developed for use by teachers and school professionals. The study by Morales-Hidalgo et al. (2017) did not provide any information on administration time. Similarly, there was no information on administration time provided for the AABC, a 57-item questionnaire developed to be completed by parents, teachers, primary caregivers or persons familiar with the individual (Özdemir & Diken, 2018). Estimating the administration time based on the 11 or 57 questions introduces subjectivity when compared to the SCQ, which has 40 questions and takes 10 minutes. The EDUTEA and AABC are emerging tools and having more comprehensive information would have aided in forming an opinion about their usefulness in Africa.

For intellectual disabilities, a total of 6 tools were identified: the HASI, LDSQ, CAIDS-Q, SIT, SCIL and QT. Two of these tools (SIT and QT) focus solely on IQ scores to determine the presence of intellectual disabilities. Moreover, the QT is rather outdated and also measures mostly verbal skills, based on old-fashioned pictures which may not be culturally relevant to African settings. For individuals not verbally able, in Africa, the QT will not be very useful. The original SIT was considered outdated and not on a par with the Wechsler scales (Kunen et al.,

1996) and was revised to address some of the concerns. However, the new SIT-R still focuses on verbal cognitive ability. In addition, as previously mentioned, other reviews of the SIT-R have mentioned problems associated with the reliability and validity of the tool. The LDSQ, meanwhile, is adult-specific, and studies that used the LDSQ had participants aged 18-years and above; 18-years is considered the legal adult age in most developed economies. The challenge posed lies in the lower age limit of 18, implying that the LDSQ cannot be used with persons younger than 18-years old.

To close the gap, the CAIDS-Q was developed, by the authors of the LDSQ. CAIDS-Q is used for 8 – 18-year-olds. For screening adolescents, as defined by age 11 – 26 years, a more encompassing single measure is required. Two screening tools met this criterion, the HASI and the SCIL. HASI can be used with persons as young as 10-years, as there are two different cut-off scores: one for those below 18-years and another for those above 18-years. Given that HASI requires some training to use it and is also used largely in the CJS and forensic services, two areas that are underrepresented in the African context, these may impact on its usefulness in the African environment. The SCIL was developed and examined with adults (18 – 63 years) and adolescents (12 – 17 years). The SCIL also incorporates test items that assess social adaptive skills in line with the current diagnostic criteria for intellectual disabilities, as per the DSM-5, but is currently only available in the Dutch language.

A combined total of 18 screening tools were reviewed for autism and intellectual disabilities. The quality of the tool's design, studies employing them, and overall evidence provided were analysed using the COSMIN Risk of Bias checklist (Tables 2 through 7). Examples of the areas analysed were the concept elicitation, clearly describing the construct of interest, target population, and context of use. Based on the results of the review using the COSMIN checklist and additional information from manuals, the overall ratings for twelve tools (SCQ, CARS, PDD-MRS, EDUTEA, AABC, DiBAS-R, DBC-ASA, CBCL, LDSQ, CAIDS-Q, HASI and SCIL) were moderate. For four tools (AQ, MARA, SIT/SIT-R, and ADI-TSS) the rating was low, and the remaining 2 tools (ASSQ, and QT) were rated as very low.

Psychometric properties

For autism, clinical samples with a previous diagnosis participated in most of the studies, leading to a focus upon discriminative validity, differentiating those with and without autism. Using

clinical data also meant that a comparison of the outcomes from the screening tools was not necessarily compared to those of an acceptable gold standard instrument. Regarding sensitivity, specificity, PPV and NPV, most of the studies reported values for specificity and sensitivity only. Psychometric properties from the studies reviewed were quite varied (Table 8). The variations could be due to the heterogeneity of the participants across age, gender, severity, or the adjustment in cut-off scores. Other factors that can impact outcomes are study methodology and proxy informants (Ehlers & Gillberg, 1993). Deriving a cut-off score that is associated with precision is part of the development of instruments; however, in some studies, these adjustments resulted in marked variations. This variability was exemplified in studies that utilised the LDSQ. The studies by McKenzie et al. (2012a, 2015) reported sensitivities of .82 and .92, respectively, while Stirk et al. (2018) reported a sensitivity of .67.

Applying the guidelines from the COSMIN checklist, the quality of studies on measurement properties and the evidence for those properties were analysed (details in Tables 2 and 3, see Appendix). The properties included content validity (this includes relevance of the items in the tool, comprehensiveness, and comprehensibility), structural validity, internal consistency, reliability, criterion validity and construct validity. In rating the content validity, expert and end-user input are considered. COSMIN ratings are based on the 'lowest score' counts, as previously mentioned, and this formed the basis for the overall rating of the studies and outcomes. Eight tools (PDD-MRS, DiBAS-R, CBCL, DBC-ASA, MARA, AABC, SCQ & EDUTEA) had moderate evidence for content validity with the remaining four rated low. There was moderate evidence for structural validity for eight tools (SCQ, CARS, CBCL, DBC-ASA, MARA, EDUTEA, DiBAS-R & PDD-MRS). AABC had high evidence for structural validity while the remaining 3 had low or very low evidence. Only 82% (18) of the studies examined criterion validity, and these were studies that used the EDUTEA, PDD-MRS, ADI-TSS, DiBAS-R, AQ-10, ASSQ, SCQ, AABC and DBC-ASA. Out of these, the PDD-MRS, EDUTEA and SCQ were rated high while the DiBAS-R, AABC and ADI-TSS were rated moderate. Evidence from the remaining three tools was inadequate, and they received ratings of low. There was enough evidence to give a rating of moderate to the ADI-TSS, EDUTEA, SCQ, and DiBAS-R for construct validity while the PDD-MRS received a rating of high. Reliability was high in the EDUTEA and moderate for SCQ, CBCL, and CARS. Internal consistency was found to be high in the EDUTEA and moderate for the SCQ, PDD-MRS, AABC, CARS, DiBAS-R and AQ-10.

Some ratings for the DBC-ASA and the CBCL were based on the manuals not on the studies. When all components of the psychometric properties are considered, the SCQ, CBCL, DBC-ASA, and PDD-MRS met most of the COSMIN criteria.

Turning to consider intellectual disabilities, out of the nineteen studies reviewed, fifteen of them which used the HASI, CAIDS-Q, LDSQ and SCIL incorporated the current DSM-5 criteria for intellectual disabilities by using both IQ and adaptive behaviour in the tools. The other four based on the SIT and QT focused on making comparisons with Full-Scale IQ as a basis for identifying participants with intellectual disabilities. Seventeen of the studies reviewed (89%) validated their results against the age-appropriate Wechsler scales, the most widely used assessment of general intellectual functioning, and often regarded as the gold standard. One study (Kunen et al., 1996) compared the SIT to the Stanford-Binet while Rotatori & Epstein (1978) focused on test-retest reliability.

All studies involving people with intellectual disabilities had evidence of explicit constructs for the development of the tools and, like the autism studies, these studies examined mainly the discriminative and predictive validities of the measures. Criterion validity was examined in all the studies with the HASI, and SCIL rated as high; while those with SIT, CAIDS-Q, LDSQ and QT were rated moderate. Evidence for construct validity was high for the HASI and CAIDS-Q and moderate for SIT, SCIL, LDSQ and QT. Internal consistency was high in the SCIL, moderate for CAIDS-Q and low for the HASI, LDSQ and QT, and very low for the SIT. The quality of evidence for content validity was moderate for the SCIL, CAIDS-Q, LDSQ and QT while very low for SIT and low for the HASI.

Regarding reliability, the HASI, QT and LDSQ were rated low while the SCIL, SIT and CAIDS-Q were moderate. Structural validity was rated high for the SCIL, moderate for CAIDS-Q and low to very low for the remaining four. Sensitivity, specificity, PPV and NPV values from the studies were also varied but generally within acceptable ranges (Glascoe 2005). Sensitivity was between .67 and 1 while specificity was between .35 and .92. Based on this review, none of the intellectual disabilities screening tools identified through this review seemed to have been used in Africa. The SCIL and CAIDS-Q were found to have better overall psychometric properties and scored better on the COSMIN checklist (Tables 3 and 7, see Appendix). Not all studies incorporated adaptive behaviour scores alongside IQ and overall, in the future, there needs to be

more of a shift from IQ testing as a measure of intellectual disabilities, to incorporating adaptive skills during screening and eventual diagnosis by using a tool that captures both.

Cultural adaptation

A key element for any of the tools selected for use within African nations is cross-cultural validity. Cross-cultural validity based on the COSMIN checklist includes the sample size, agreement between the original and translated versions, use with different populations, diagnoses and ethnicities. For example, a Spanish version compared to an English version, or Dutch participants compared to German participants or adults to adolescents.

All the screening tools identified through the review were used for both male and female participants. Comparisons were made between those with and without autism or intellectual disabilities Concerning the use of different versions of each tool, the AQ in English was used only in the UK, the English ASSQ in the USA and the Swedish version in Sweden. DiBAS-R which is in German was used in Germany, MARA in the USA, ADI-TSS Spanish version was developed and used in Argentina, PDD-MRS in the Netherlands where it originated as well as the Spanish version used in Spain, and the EDUTEA in Spain where it was developed. Four different versions of the CBCL (English version completed by 60% of the participants, the Chinese version 30%, Malay 8%, Tamil 2%) were used by Ooi et al. (2011), CARS in the USA and Spain while the DBC-ASA was used only in the USA. The SCQ was used in the UK, Qatar, Australia, and USA. The AABC was used only in Turkey. Out of the 12 screening tools for autism, the SCQ was used across a wider age range, across more disabilities, and comorbidities (Ung et al., 2016). The validity of the SCQ has also been examined in a small sample of children aged between 2.5 and 14-years in a South African study (Bozalek (2013). When all assessment criteria for cross-cultural validity were examined, the overall rating for the autism tools was as follows: very low for the ADI-TSS, AQ, and MARA; low for the ASSQ, DBC-ASA, AABC and DiBAS-R; medium for EDUTEA, PDD-MRS, CBCL, CARS, and SCQ.

Out of the 19 studies reviewed for the intellectual disabilities screening tools, five studies used between groups designs, including samples of people without intellectual disabilities, six were within-subject designs, while the remaining eight were cross-sectional designs. Utilising the tool with different groups is a criterion for cross-cultural validity in the COSMIN, so studies that have not demonstrated this adequately were rated low in that area. HASI was used in two within-

subject studies and five cross-sectional studies. SIT was used in three within-subject studies while the QT was used in one within-subject study. The LDSQ was employed in two between-subject studies, the CAIDS-Q in 3 between-subject and one cross-sectional studies while the SCIL was used in two cross-sectional studies. The HASI was used in 4 different countries and languages: Norway, UK, Australia and Belgium. LDSQ was used in the UK and Scotland while the CAIDS-Q and QT were used in the UK and USA, respectively. The SCIL was used in Norway and the Netherlands. Putting together all the criteria for evaluating cross-cultural validity, the overall rating for the tools was moderate for CAIDS-Q and HASI, high for SCIL, low for the LDSQ and very low for both the SIT and QT.

Finally, given that one of the aims of this review was to consider the appropriateness of using these tools across a range of cultures, it is important to note that there are diverse cultures in Africa. These include a variety of spoken languages, beliefs and behaviours; therefore, whichever tools are identified through this review will require additional contextual adaptation and may perhaps benefit from further ethnological research.

Limitations

There are limitations to this review. By limiting the search to studies in English only, it is possible that some studies with adolescents, and potentially other tools, may have been missed. This in turn may limit the generalisability of the findings of this review, as there are some African countries whose official languages are not English. Manuals for some of the screening tools (seven in total) identified were not readily accessible. This meant that some information on validation reported in the studies could not be compared. Additionally, some of the administration and training requirements could not be examined in detail.

Conclusion

There are two main challenges. The first relates to cultural adaptations and use of the tools outside of the development environment. Whichever tool is identified for use in Africa, it must be sensitive to local differences and expression. The language of the tool must be simple enough to understand, allowing for ease with translation or substitution where required. Validation of selected tools will require time, expertise and financial resources as determining the psychometric properties in a nouvelle environment requires capacity. As such, the less complex

the tool is, the easier it may be to assemble the required resources. These challenges are not to suggest the screening tools developed in the West are irrelevant to Africa or LMICs, but that careful research and translational work may need to be done to ensure that a tool can be used successfully with people from other countries and cultures. A second challenge is that the tools selected for use with the adolescent and young adult population need to apply to a wide age range while remaining flexible and sensitive. Finally, the limited number of studies involving adolescents identified through this review has presented challenges (as in Hirota et al., 2018). Without a large body of knowledge about adolescents and continental Africa, particularly, the choice of tools is limited.

Developing and validating a continent-specific or country-specific tool for screening autism or intellectual disabilities will take considerable time, effort and resources. Such resources as time, training and personnel may not be readily available. Given the socio-economic and political climate of most African countries, the process could place a considerable financial burden on the economies. In summary, of the 18 tools (6 for intellectual disabilities and 12 for autism) identified through the review, except for the SCQ, none had been utilised in Africa. The SCQ was designed to be used with a wide age range, 4-years and above and has two versions (current and lifetime) which makes it a good fit for use with adolescents. The SCIL, meanwhile, was validated for adolescents and adults and includes test questions for intellectual abilities as well as social adaptive skills. The broad age range and inclusion of DSM-5 items places it above the other tools reviewed. Additionally, any tool that requires training and more than 20-minutes of administration time will add to the burden. Thus, to begin the process of validating the screening tools for autism and intellectual disabilities in African adolescents, two tools seem particularly appropriate from the review. These are the SCQ for autism and the SCIL for intellectual disabilities.

Financial Disclosure

None

Funding Source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors do not have any conflict of interest to declare.

References

Achenbach, T. M., & Rescorla, L. A. (2001). Manual for the ASEBA school-age forms & profiles. Burlington, VT: University of Vermont. *Research Center for Children, Youth, & Families*.

Al Mamun, K. A., Bardhan, S., Ullah, M. A., Anagnostou, E., Brian, J., Akhter, S., & Rabbani, M. G. (2016). Smart autism - a mobile, interactive and integrated framework for screening and confirmation of autism. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference, 2016*, 5989-5992.

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(2), 202-212.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5 (Fifth Ed.).* Arlington, Virginia. American Psychiatric Publishing.

Ammons, R. B., & Ammons, C. H. (1962). The quick test (QT): provisional manual. *Psychological Reports*, 11(1), 111-161.

Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The Autism Spectrum Quotient (AQ)-adolescent version. *Journal of Autism & Developmental Disorders*, 36(3), 343-350.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism & Developmental Disorders*, 31(1), 5-17.

Barton, M. L., Dumont-Mathieu, T., & Fein, D. (2012). Screening young children for autism spectrum disorders in primary practice. Journal of Autism and Developmental Disorders, 42(6), 1165-1174.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism Screening Questionnaire: Diagnostic validity. *British Journal of Psychiatry*, 175(5), 444-451.

Blackwell, T. L., & Madere, L. N. (2005). Slosson Intelligence Test-Revised. *Rehabilitation Counseling Bulletin*, 48(3), 183.

Booth, T., Murray, A. L., McKenzie, K., Kuenssberg, R., O'Donnell, M., & Burnett, H. (2013). Brief report: An evaluation of the AQ-10 as a brief screening instrument for ASD in adults. *Journal of Autism and Developmental Disorders, 43*(12), 2997-3000.

Bozalek, F. (2013). Autism Screening in Children: Using the Social Communication Questionnaire in a Western Cape Population. University of Cape Town

Braatveit, K. J., Torsheim, T., & Hove, O. (2018). Screening for intellectual disabilities: A validation of the Hayes Ability Screening Index for in-patients with substance use disorder. *Nordic Journal of Psychiatry*, 72(5), 387-392.

Breidbord, J., & Croudace, T. J. (2013). Reliability generalization for Childhood Autism Rating Scale. *Journal of autism and developmental disorders, 43*(12), 2855-2865.

Brereton, A. V., Tonge, B. J., Mackinnon, A. J., & Einfeld, S. L. (2002). Screening young people for autism with the Developmental Behavior Checklist. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(11), 1369-1375.

Brooks, W. T., & Benson, B. A. (2013). The validity of the Social Communication Questionnaire in adults with intellectual disability. *Research in Autism Spectrum Disorders, 7*(2), 247-255.

Campbell, C. A., & Ashmore, R. J. (1995). Test Review: The Slosson Intelligence Test-Revised (SIT-R). *Measurement and Evaluation in Counseling and Development*, 28(2), 116-18.

Cederberg, C. D., Gann, L. C., Foley-Nicpon, M., & Sussman, Z. (2018). ASD screening measures for high-ability youth with ASD: Examining the ASSQ and SRS. *Gifted Child Quarterly*, *62*(2), 220-229.

Charman, T., Baird, G., Simonoff, E., Loucas, T., Chandler, S., Meldrum, D., & Pickles, A. (2007). Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *British Journal of Psychiatry*, 191(6).

Cochrane, A., & Holland, W. (1971). Validation of screening procedures. *British Medical Bulletin*, 27(1), 3-8.

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook Jr, 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.

Deb, S., Dhaliwal, A. J., & Roy, M. (2009). The Usefulness of the DBC-ASA as a screening instrument for autism in children with intellectual disabilities: A pilot study. *Journal of Applied Research in Intellectual Disabilities*, *22*(5), 498-501.

DiLalla, D. L., & Rogers, S. J. (1994). Domains of the Childhood Autism Rating Scale: Relevance for diagnosis and treatment. *Journal of Autism and Developmental Disorders*, *24*(2), 115-128.

Duda, M., Daniels, J., & Wall, D. (2016). Clinical evaluation of a novel and Mobile Autism Risk Assessment. *Journal of Autism & Developmental Disorders*, *46*(6), 1953-1961.

Durkin, M. (2001). Measurement of childhood disabilities in population studies. Paper presented at the *International Seminar on Measurement of Disability, New York, USA.*

Ehlers, S., & Gillberg, C. (1993). The epidemiology of Asperger Syndrome: A total population study. *Child Psychology & Psychiatry & Allied Disciplines*, 34(8), 1327-1350.

Ehlers, S., Gillberg, C., & Wing, L. (1999). A screening questionnaire for Asperger Syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of Autism and Developmental Disorders*, 29(2), 129-141.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology, 38*(3), 439-450.

Emerson, E. (2012). Deprivation, ethnicity and the prevalence of intellectual and developmental disabilities. *Journal of Epidemiology and Community Health*, 66(3), 218-224.

Ford, G., Andrews, R., Booth, A., Dibdin, J., Hardingham, S., & Kelly, T. P. (2008). Screening for learning disability in an adolescent forensic population. *The Journal of Forensic Psychiatry & Psychology*, 19(3), 371-381.

Franz, L., Chambers, N., von Isenburg, M., & de Vries, P. J. (2017). Autism spectrum disorder in sub-Saharan Africa: A comprehensive scoping review. *Autism Research*, 10(5), 723-749.

Gau, S. S., Lee, C., Lai, M., Chiu, Y., Huang, Y., Kao, J., & Wu, Y. (2011). Psychometric properties of the Chinese version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders*, 5(2), 809-818.

Geijsen, K., Kop, N., & de Ruiter, C. (2018). Screening for intellectual disability in Dutch police suspects. *Journal of Investigative Psychology and Offender Profiling*, 15(2), 200-214.

Gladstone, M., Lancaster, G. A., Umar, E., Nyirenda, M., Kayira, E., van den Broek, N. R., & Smyth, R. L. (2010). The Malawi Developmental Assessment Tool (MDAT): The creation, validation, and reliability of a tool to assess child development in rural African settings. *PLoS Medicine*, 7(5).

Gladstone, M., Mallewa, M., Jalloh, A. A., Voskuijl, W., Postels, D., Groce, N., Kerac, M. & Molyneux, E. (2014). Assessment of neurodisability and malnutrition in children in Africa. *Seminars in Pediatric Neurology*, 21(1), 50-57.

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(3), 173-179.

Guevara, J. P., Gerdes, M., Localio, R., Huang, Y. V., Pinto-Martin, J., Minkovitz, C. S., Pati, S. (2013). Effectiveness of developmental screening in an urban setting. *Pediatrics*, 131(1), 30.

Guo, Y., Tang, Y., Rice, C., Lee, L., Wang, Y., & Cubells, J. F. (2011). Validation of the Autism Spectrum Screening Questionnaire, Mandarin Chinese version (CH-ASSQ) in Beijing, China. *Autism, 15*(6), 713-727

Gura, G. F., Champagne, M. T., & Blood-Siegfried, J. (2011). Autism spectrum disorder screening in primary care. *Journal of Developmental & Behavioral Pediatrics, 32*(1), 48-51.

Hasegawa, J., Ito, Y. M., & Yamauchi, T. (2017). Development of a screening tool to predict malnutrition among children under two years old in Zambia. *Global Health Action*, 10(1), 1339981.

Hayes, S. C. (2002). Early intervention or early incarceration? Using a screening test for intellectual disability in the criminal justice system. *Journal of Applied Research in Intellectual Disabilities*, 15(2), 120-128.

Heinrich, M., Böhm, J., & Sappok, T. (2018). Diagnosing autism in adults with intellectual disability: Validation of the DiBAS-R in an independent sample. *Journal of Autism & Developmental Disorders, 48*(2), 341-350.

Hirota, T., So, R., Kim, Y. S., Leventhal, B., & Epstein, R. A. (2018). A systematic review of screening tools in non-young children and adults for autism spectrum disorder. *Research in Developmental Disabilities*, 80, 1-12.

Jinabhai, C. C., Taylor, M., Rangongo, M. F., Mkhize, N. J., Anderson, S., Pillay, B. J., & Sullivan, K. R. (2004). Investigating the mental abilities of rural Zulu primary school children in South Africa. *Ethnicity & health*, 9(1), 17-36.

Kaal, H.L., Nijman, H.L.I. and Moonen, X.M.H. (2015), "Identifying offenders with an intellectual disability in detention in The Netherlands". *Journal of Intellectual Disabilities and Offending Behaviour*, Vol. 6 No. 2, pp. 94-101.

Kakooza-Mwesige, A., Ssebyala, K., Karamagi, C., Kiguli, S., Smith, K., Anderson, M. C., Smith, D. (2014). Adaptation of the "Ten Questions" to screen for autism and other neurodevelopmental disorders in Uganda. *Autism*, 18(4), 447-457.

Knox, J., Arpadi, S. M., Kauchali, S., Craib, M., Kvalsvig, J. D., Taylor, M., Bah, F., Mellins, C., & Davidson, L. L. (2018). Screening for developmental disabilities in HIV positive and HIV negative children in South Africa: Results from the Asenze study. PloS One, 13(7)

Kopp, S., & Gillberg, C. (2011). The Autism Spectrum Screening Questionnaire (ASSQ)-Revised Extended Version (ASSQ-REV): An instrument for better capturing the autism phenotype in girls? A preliminary study involving 191 clinical cases and community controls. *Research in Developmental Disabilities*, 32(6), 2875-2888.

Koura, K. G., Boivin, M. J., Davidson, L. L., Ouédraogo, S., Zoumenou, R., Alao, M. J., Garcia, A., Massougbodji, A., Cot, M., & Bodeau-Livinec, F. (2013). Usefulness of child development assessments for low-resource settings in francophone Africa. *Journal of Developmental and Behavioral Pediatrics*: *JDBP*, 34(7), 486–493.

Kraijer, D. W. (1990). *AVZ, Autisme- en Verwante kontaktstoornissenschaal voor Zwakzinnigen. Handleiding (Autism and related contact disorders scale for the mentally retarded. Manual)*. Lisse, The Netherlands: Swets and Zeitlinger.

Kraijer, D. (1994). *AVZ-R, Autisme- en Verwante stoornissenschaal voor Zwakzinnigen-Revisie. Handleiding. Sterk herziene en uitgebreide uitgave (Autism and related disorders scale for the mentally retarded-Revision. Manual. Second revised edition)* Lisse, The Netherlands: Swets and Zeitlinger

Kraijer, D., & de Bildt, A. (2005). The PDD-MRS: An instrument for identification of autism spectrum disorders in persons with mental retardation. *Journal of Autism and Developmental Disorders*, 35(4), 499-513.

Krug, D. A., Arick, J., & Almond, P. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry*, *21*(3), 221-229.

Kunen, S., Overstreet, S., & Salles, C. (1996). Concurrent validity study of the Slosson Intelligence Test-Revised in mental retardation testing. *Mental Retardation, 34*(6), 380-386.

Lesinskiene, S. (2000). Vilniaus miesto vaiku autizmas, vilniaus universitetas, daktaro disetacijos santrauka, biomedicinos mokslai, medicina 07B, *Psichiatrija* B650. Lithuania: Vilnius,

Limbos, M. M., & Joyce, D. P. (2011). Comparison of the ASQ and PEDS in screening for developmental delay in children presenting for primary care. *Journal of Developmental & Behavioral Pediatrics*, 32(7), 499-511.

Luckasson, R., & Schalock, R. L. (2013). What's at stake in the lives of people with intellectual disability? Part II: Recommendations for naming, defining, diagnosing, classifying, and planning supports. *Intellectual and Developmental Disabilities*, 51(2), 94-101.

Maxim, L. D., Niebo, R., & Utell, M. J. (2014). Screening tests: A review with examples. *Inhalation Toxicology*, 26(13), 811-828.

Mazefsky, C. A., Anderson, R., Conner, C. M., & Minshew, N. (2011). Child Behavior Checklist scores for school-aged children with autism: Preliminary evidence of patterns suggesting the need for referral. *Journal of Psychopathology and Behavioral Assessment*, 33(1), 31-37.

McConachie, H., Parr, J. R., Glod, M., Hanratty, J., Livingstone, N., Oono, I. P., et al. (2015). Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technology Assessment*, 19(41), 1–506

McKenzie, K. and Paxton, D. (2006). *Learning Disability Screening Questionnaire*. GCM Records, Edinburgh.

McKenzie, K. and Paxton D. (2012). *Child and Adolescent Intellectual Disability Screening Questionnaire*. GCM Records, Edinburgh.

McKenzie, K., Michie, A., Murray, A., & Hales, C. (2012a). Screening for offenders with an intellectual disability: the validity of the Learning Disability Screening Questionnaire. Research in Developmental Disabilities, 33(3), 791-795.

McKenzie, K., Murray, G., Murray, A., Delahunty, L., Hutton, L., Murray, K., & O'Hare, A. (2019). Child and Adolescent Intellectual Disability Screening Questionnaire to identify children with intellectual disability. *Developmental Medicine & Child Neurology*, 61(4), 444-450.

McKenzie, K., Paxton, D., Michie, A., Murray, G., Murray, A., & Curtis, J. (2012b). Screening with young offenders with an intellectual disability. *Journal of Forensic Psychiatry & Psychology*, 23(5-6), 676-688.

McKenzie, K., Paxton, D., Murray, G., Milanesi, P., & Murray, A. L. (2012c). The evaluation of a screening tool for children with an intellectual disability: The Child and Adolescent Intellectual Disability Screening Questionnaire. *Research in Developmental Disabilities*, 33(4), 1068-1075.

Mesibov, G. B., Schopler, E., Schaffer, B., & Michal, N. (1989). Use of the Childhood Autism Rating Scale with autistic adolescents and adults. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*(4), 538-541.

Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018a). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1171-1179.

Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2018b). COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). *User manual*. *https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf*

Morales-Hidalgo, P., Hernández-Martínez, C., Voltas, N., & Canals, J. (2017). EDUTEA: A DSM-5 teacher screening questionnaire for autism spectrum disorder and social pragmatic

communication disorder. *International Journal of Clinical Health & Psychology, 17*(3), 269-281.

Mouti, A., Dryer, R., & Kohn, M. (2019). Differentiating autism spectrum disorder from ADHD using the Social Communication Questionnaire. *Journal of Attention Disorders, 23*(8), 828-837.

Mung'ala-Odera, V., Meehan, R., Njuguna, P., Mturi, N., Alcock, K., Carter, J. A., & Newton, C. R. (2004). Validity and reliability of the 'ten questions' questionnaire for detecting moderate to severe neurological impairment in children aged 6-9 years in rural Kenya. *Neuroepidemiology*, 23(1-2), 67-72.

Murphy, G. H., Gardner, J., & Freeman, M. J. (2017). Screening prisoners for intellectual disabilities in three English prisons. *Journal of Applied Research in Intellectual Disabilities*, 30(1), 198-204.

Nijman, H., Kaal, H., van Scheppingen, L., & Moonen, X. (2018). Development and testing of a Screener for Intelligence and Learning Disabilities (SCIL). *Journal of Applied Research in Intellectual Disabilities, 31*(1), e59-e67.

Özdemir, O., & Diken, I. H. (2019). Reliability and Validity Studies of the Adapted Autism Behaviour Checklist in Turkey. *Journal of Developmental and Physical Disabilities*, *31*(3), 359-376.

Ooi, Y. P., Rescorla, L., Ang, R. P., Woo, B., & Fung, D. S. S. (2011). Identification of autism spectrum disorders using the Child Behavior Checklist in Singapore. *Journal of Autism and Developmental Disorders, 41*(9), 1147-1156.

Olusanya, B. O., & Okolo, A. A. (2006). Revisiting the Ten Questions questionnaire for developing countries. *International Journal of Epidemiology*, 35(4), 1103-1103.

Oshodi, Y. O., Olagunju, A. T., Oyelohunnu, M. A., Campbell, E. A., Umeh, C. S., Aina, O. F., Adeyemi, J. D. (2017). Autism spectrum disorder in a community-based sample with

neurodevelopmental problems in Lagos, Nigeria. *Journal of Public Health in Africa, 7*(2), 559; 559-559.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et. Al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71.

Paxton, D., McKenzie, K., & Murray, G. (2008). Putting screening tools to the test. *Learning Disability Practice, 11*(8), 14-19.

Posserud, M. B., Lundervold, A. J., & Gillberg, C. (2006). Autistic features in a total population of 7–9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire). *Journal of Child Psychology and Psychiatry, 47*(2), 167-175.

Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C.W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research, 27*(5), 1147-1157.

Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., Döpfner, M. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders, 15*(3), 130-142.

Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism, 12*(5), 537-556.

Rotatori, A. F., & Epstein, M. (1978). The Slosson Intelligence Test as a quick screening test of mental ability with profoundly and severely retarded children. *Psychological Reports*, 42(3_suppl), 1117-1118.

Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism Diagnostic Interview-Revised. Los Angeles, CA: Western Psychological Services.

Saloojee, G., Phohole, M., Saloojee, H., & IJsselmuiden, C. (2007). Unmet health, welfare and educational needs of disabled children in an impoverished South African peri-urban township. *Child: Care, Health and Development*, 33(3), 230-235.

Sappok, T., Gaul, I., Bergmann, T., Dziobek, I., Bölte, S., Diefenbacher, A., & Heinrich, M. (2014a). The Diagnostic Behavioral Assessment for Autism Spectrum Disorder—Revised: A screening instrument for adults with intellectual disability suspected of autism spectrum disorders. *Research in Autism Spectrum Disorders, 8*(4), 362-375.

Sappok, T., Gaul, I., Dziobek, I., Bölte, S., Diefenbacher, A., & Bergmann, T. (2014b). Der Diagnostische Beobachtungsbogen für Autismus Spektrumstörungen (DiBAS): Ein Screeninginstrument für Erwachsene mit Intelligenzminderung bei Autismusverdacht. *Psychiatrische Praxis*, 41, 1–7.

Sawyer, S. M., Afifi, R. A., Bearinger, L. H., Blakemore, S., Dick, B., Ezeh, A. C., & Patton, G. C. (2012). Adolescence: A foundation for future health. *The Lancet*, 379(9826), 1630-1640.

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. *The Lancet Child & Adolescent Health*, 2(3), 223-228.

Sawyer, R. N., & Whitten, J. R. (1972). Concurrent validity of the Quick Test. Psychological reports, 30(1), 64-66.

Schalock, R. L. & Luckasson, R. A. (2013) What's at stake in the lives of people with intellectual disability? Part I: The power of naming, defining, diagnosing, classifying, and planning supports. *Intellectual and Developmental Disabilities*, 51(2), 86-93.

Scherzer, A. L., Chhagan, M., Kauchali, S., & Susser, E. (2012). Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. *Developmental Medicine & Child Neurology*, 54(12), 1079-1084.

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.

Snow, A. (2013). Social Communication Questionnaire. In F. R. Volkmar (Ed.), Encyclopedia of autism spectrum disorders (pp. 2893-2895). New York, NY: Springer New York.

Søndenaa, E., Bjørgen, T. G., & Nøttestad, J. A. (2007). Validation of the Norwegian version of Hayes Ability Screening Index for mental retardation. *Psychological reports*, *101*(3_suppl), 1023-1030.

Søndenaa, E., Nygård, Ø., Nøttestad, J. A., & Linaker, O. M. (2011). Validation and adaptation of the Norwegian version of Hayes Ability Screening Index for intellectual difficulties in a psychiatric sample. *Nordic Journal of Psychiatry*, *65*(1), 47-51.

Søndenaa, E., Rasmussen, K., Palmstierna, T., & Nøttestad, J. (2008). The prevalence and nature of intellectual disability in Norwegian prisons. *Journal of Intellectual Disability Research, 52*(12), 1129-1137.

Soto, S., Linas, K., Jacobstein, D., Biel, M., Migdal, T., & Anthony, B. J. (2015). A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism*, 19(6), 646-661.

Stein, Z., Durkin, M., & Belmont, L. (1986). "Serious" mental retardation in developing countries: An epidemiologic approach. *Annals of the New York Academy of Sciences, 477*(1), 8-21

Steiner, A. M., Goldsmith, T. R., Snow, A. V., & Chawarska, K. (2012). Practitioner's guide to assessment of autism spectrum disorders in infants and toddlers. *Journal of Autism and Developmental Disorders*, 42(6), 1183-1196.

Stirk, S., Field, B., & Black, J. (2018). An independent investigation of the utility of the Learning Disability Screening Questionnaire (LDSQ) within a community learning disability team. *Journal of Applied Research in Intellectual Disabilities, 31*(2), e223-e228.

Suhail, K., & Zafar, F. (2008). Prevalence of autism in special education schools of Lahore. Pakistan *Journal of Psychological Research*, 23(2)

Swinkels, S. H., Dietz, C., van Daalen, E., Kerkhof, I. H., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum in children aged 14 to 15 months. I: the development

of the Early Screening of Autistic Traits Questionnaire (ESAT). *Journal of Autism and Developmental Disorders*, 36(6), 723-732.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research, 27*(5), 1159-1170.

To, W. T., Vanheule, S., Vanderplasschen, W., Audenaert, K., & Vandevelde, S. (2015). Screening for intellectual disability in persons with a substance abuse problem: Exploring the validity of the Hayes Ability Screening Index in a Dutch-speaking sample. *Research in Developmental Disabilities, 36*, 498-504.

Tomlinson, M., Yasamy, M. T., Emerson, E., Officer, A., Richler, D., & Saxena, S. (2014). Setting global research priorities for developmental disabilities, including intellectual disabilities and autism. *Journal of Intellectual Disability Research,* 58(12), 1121-1130.

Trivedi, A. (1977). A comparison of three intelligence tests for the assessment of mental retardation. *Journal of Mental Deficiency Research*, *21*(4), 289-297.

Ung, D., Johnco, C., McBride, N. M., Howie, F., Scalli, L., & Storch, E. A. (2016). Optimizing the screening of autism spectrum disorders in outpatient clinics: An examination of the social communication questionnaire-lifetime. *Research in Autism Spectrum Disorders, 27*, 21-28.

United Nations. (2014). Country classification: Data sources, country classifications and aggregation methodology. *World Economic Situation and Prospects 2014*, 143-150.

United Nations Department of Economic and Social Affairs. (2021). *World Economic Situation and Prospects 2021*. United Nations.

Van der Linde, J., Swanepoel, D. W., Glascoe, F. P., Louw, E. M., & Vinck, B. (2015). Developmental screening in South Africa: Comparing the national developmental checklist to a standardized tool. *African Health Sciences*, 15(1), 188-196.

Vawda, N., Milburn, N. G., Steyn, R., & Zhang, M. (2017). The development of a screening tool for the early identification of risk for suicidal behavior among students in a developing country. Psychological Trauma: Theory, *Research, Practice, and Policy*, 9(3), 267.

Vrancic, D., Nanclares, V., Soares, D., Kulesz, A., Mordzinski, C., Plebst, C., & Starkstein, S. (2002). Sensitivity and specificity of the Autism Diagnostic Inventory-Telephone Screening in Spanish. *Journal of Autism & Developmental Disorders, 32*(4), 313-320

Wall, D. P., Dally, R., Luyster, R., Jung, J. Y., & Deluca, T. F. (2012). Use of artificial intelligence to shorten the behavioral diagnosis of autism. *PLoS ONE*, 7(8).

Webb, E., Morey, J., Thompsen, W., Butler, C., Barber, M., & Fraser, W. (2003). Prevalence of autistic spectrum disorder in children attending mainstream schools in a Welsh education authority. *Developmental Medicine & Child Neurology*, 45(6), 377-384.

Westerlund, M., & Sundelin, C. (2000). Screening for developmental language disability in 3-year-old children. experiences from a field study in a Swedish municipality. *Child: Care, Health and Development: Original Articles*, 26(2), 91-110.

Wild, D., Furtado, T., & Angalakuditi, M. (2012). The translation and cultural adaptation of the Child Behavior Checklist for use in Israel (Hebrew), Korea, the US (Spanish), India (Malayalam and Kannada), and Spain. *Psychology Research and Behavior Management*, 5, 51-56.

World Bank. (2020). *World Bank country and lending groups*. Retrieved from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519

Appendix 21 – Focus Group – Submitted Manuscript

Link to the article –

Using the consensus group method to select the best screening tools for autism and intellectual disability for use with Nigerian adolescents

Eziafakaku Uchechukwu Nwokolo and Glynis H Murphy

University of Kent

Anne-Marie Mensink

Norfolk Community Health and Care NHS Trust

Xavier Moonen

University of Amsterdam

Peter E Langdon

University of Warwick and Coventry and Warwickshire Partnership NHS Trust

Author Note

Eziafakaku Uchechukwu Nwokolo, Tizard Centre, University of Kent; Glynis H Murphy, Tizard Centre, University of Kent; Anne-Marie Mensink, Norfolk Community Health and Care NHS Trust; Xavier Moonen, University of Amsterdam; Peter E Langdon, Centre for Educational Development, Appraisal and Research (CEDAR), and Centre for Mental Health and Wellbeing Research, University of Warwick and Brooklands Hospital, Coventry and Warwickshire Partnership NHS Trust.

Correspondence concerning this article should be addressed to Eziafakaku Nwokolo, Tizard Centre, University of Kent, Canterbury, CT2 7NF.

Email: eun5@kent.ac.uk

Abstract

The Nominal Group Technique (NGT) was used with a purposive group of professionals, parents, and laypersons to select and adapt existing screening tools for autism spectrum disorder and intellectual disability for use with older children and adolescents in Nigeria. We identified four screening tools through a previously completed systematic review, two for autism spectrum disorder (ASD) and two for intellectual disability (ID). Both appeared appropriate for validation for use within African nations. Consultation exercises were undertaken to determine which measures may be most suitable for use within Nigeria. The group narrowed the tools down to one each for ASD and ID. The selected tools were examined for cultural relevance by the group. Following the discussions, items were either (1) accepted in the original form or (2) more culturally appropriate examples chosen if at least 75% of participants agreed. The minimum agreement on all ASD and ID measures items was 84% and indicated the measures had face and content validity for use within Nigeria.

Keywords: adolescent, screening/diagnosis, autism, intellectual disabilities, Nigeria, Africa

Introduction

Well-developed screening tools for autism spectrum disorder or intellectual disability are readily available for younger children in the West and high-income countries (McKenzie, Paxton, Murray, Milanesi & Murray, 2012; Young, 2007; Robins, Fein, Barton & Green, 2001). Also, adaptations of existing screening tools for younger children have been conducted in other countries such as Australia, Singapore, Spain, and Argentina (Nah, Young, Brewer, & Berlingeri, 2014; Canal-Bedia et al., 2011; Cuesta-Gómez, Andrea Manzone, & Posada-De-La-Paz, 2016; García-Primo et al., 2014). However, similar tools are not readily available for older children and adolescents (11 to 26-year-olds), especially in low- to middle-income countries. Very little work has been done in Africa and other low to middle-income economies regarding adapting existing tools for screening for either autism spectrum disorder or intellectual disability.

Screening for autism and intellectual disability remains a challenge in low to middle-income countries such as Nigeria due to the absence of adequate tools and other factors such as denial and low level of awareness among parents and professionals about these disabilities. Limited financial and human resources significantly contribute to the lack of adequate tools. To begin addressing this challenge in countries such as Nigeria, the adaptation of existing screening tools should be considered. Adapting existing tools is the most common and fastest approach to creating usable screening tools for countries with limited resources or expertise. However, concerns have been raised about the feasibility of employing adapted tools for screening across cultural groups (Soto et al., 2015). One way of addressing these concerns is to follow clearly defined methodologies such as those stipulated by the International Test Commission (International Test Commission, 2017). Pertinent aspects of the methodology include examining the tool's content validity, cultural validity, and language by engaging the relevant experts. Cultural validity assesses whether constructs and language initially generated in a single culture are appropriate, relevant, applicable, equivalent, and meaningful in another culture (Matsumoto & Yoo, 2006; Beaton et al., 2000). Content validity, which ensures that the items in a screening tool represent all relevant aspects of a given construct, is one of the essential psychometric properties of a screening tool (Mokkink et al., 2018, Prinsen et al., 2018; Terwee et al., 2018a). Cultural and content validity outside the original environment is usually examined by a group of experts in the environment concerned, in this case, Nigeria.

Following the completion of a recent systematic review (Nwokolo, Langdon & Murphy, submitted), twelve (12) screening tools for autism spectrum disorder (ASD) and six (6) for intellectual disability (ID) were identified. Of these, four tools were chosen (two for ASD and ID) for use within the current study

based on the cross-cultural validity and overall quality ratings of studies developing the tools. The tools for ID were (a) The adolescent version of the Screener for Intelligence and Learning Disabilities (SCIL 14-17) (Nijman et al., 2018), a standardised 14-item questionnaire developed and used for adolescents in the Netherlands, and (b) Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q) (McKenzie and Paxton, 2012), a short 7-item screening questionnaire, developed in the UK. The SCIL 14-17 was originally in Dutch and, as part of this study, translated to English, while the CAIDS-Q was in English. For screening ASD, the measures selected were the (a) Social Communication Questionnaire (SCQ) (Rutter et al., 2003) and (b) Autism Spectrum Quotient (AQ-10), the adolescent version (Allison et al., 2012). Selecting suitable and culturally sensitive measures was crucial to adapting any screening tools for use in Nigeria. Thus, a consensus group of the relevant professionals and lay people resident in Nigeria were recruited for the study. The aims of the study were to consider the face, content, and cultural validities of our chosen screening tools and make recommended adaptations for use with Nigerian adolescents using a consensus group methodology.

Method

Consensus method and Choice of Experts

The consensus group methodology was chosen due to its extensive use in studies for similar decision-making (Humphrey-Murto et al., 2017; International Test Commission, 2017). The process is based on the notion that valid, accurate and reliable evaluation is best achieved by consulting a team of experts and stakeholders. Achieving accurate and reliable assessment is assumed to be achievable through the group (Humphrey-Murto et al., 2017), and consensus methods have been used in education for curriculum development (O'Neil & Jackson, 1983), in medical and health research (Humphrey-Murto et al., 2017; Van de Ven & Delbecq, 1972). Several studies support the use of the consensus group methods in developing items for measurement tools, developing clinical guidelines, and deciding on components of new or revised curricula (Van de Ven & Delbecq, 1972; Murphy et al., 1998; Humphrey-Murto et al., 2017). Another reason for using consensus methods is that they control for possible researcher bias, and thus an appropriate and systematic process must be employed to select the best option, outcome, or measure. The consensus group method is such a technique (Delbecq, 1967; Hutchings et al., 2010 & 2012). Consensus methods are considered broadly qualitative and a systematic means for determining and developing consensus. The goal is to establish how well experts and stakeholders agree on an issue through consultation and accepting the group agreement (Tammela, 2013). This method also allows for a consideration of the cultural relevance of each measure and for associated adaptations to address any issues.

Two main techniques are used for consensus group meetings: the Nominal Group Technique (NGT) or the Delphi method. Each method raises questions, solutions proffered, and responses are ranked and agreed upon. Each of these methods has its strengths and weaknesses. Although the Delphi method is used often for the development of initial research questions and involves many anonymous participants, the Delphi method limits discussions. In contrast, the NGT involves a smaller number of participants and allows for face-to-face discussions and debates. Given that we chose existing tools and aimed to ascertain the cultural relevance, the NGT was selected as it provides for such discussion.

The nominal group technique was used to review, evaluate, and consider our screening tools' face, content, and cultural validities within a Nigerian context and make any relevant adaptations. The technique has also been applied for problem-solving and planning (Delbecq & Van de Ven, 1971), team decision-making (Bartunek & Murninghan, 1984) and as a research instrument (Van de Ven & Delbecq, 1972). NGT is a semi-quantitative, highly structured and facilitated group-based decision-making process. The process is deemed an excellent form of brainstorming with member-to-member discussions. Facilitation of discussions allows for and encourages the active participation of all members. Facilitating the group's discussion is considered beneficial as it overrides the potential of an individual member's dominance of the discussions (McMillan et al., 2016; Murphy et al., 1998). The face-to-face interactive nature of the NGT usually involves 5 – 12 participants (O'Neil & Jackson, 1983; Tammela, 2013; Humphrey-Murto et al., 2017). Where the group size is greater than this, the suggestion is that sub-groups of 8 – 10 members can be formed (O'Neil & Jackson, 1983). Delbecq and Van de Ven (1971) outlined the process for implementing the nominal group model. Although there exists a set of guidelines and a structure for using the NGT, in practice, however, the techniques have been varied based on the project or user requirements (McMillan et al., 2016; Murphy et al., 1998). Such variations may be due to the participants' time, research goals, or consensus. At other times, the required variation may be an adaptation to the stages, such as reviewing an existing protocol, measure or where the population is culturally or linguistically diverse (McMillan et al., 2016). A modified NGT was used to select and decide which autism and intellectual disability screening tools would be used for the validation study.

Experts, in the context of the NGT, are individuals who are knowledgeable about the subject matter. Given this objective, the recruitment of experts was purposive to include members from the relevant professions with professional experience and knowledge of the relevant population. For existing measures, content validity is evaluated by systematically asking professionals and users about the items' comprehensiveness, relevance and comprehensibility (Terwee et al., 2018b). A parent and layperson were included to assess the screening tools' comprehensibility. At the same time, comprehensiveness and relevance were evaluated by the professionals (Terwee et al., 2018b).

Participants

The group consisted of eight participants (60% were female, and 40% were male): a psychologist, a psychiatrist, a teacher, a paediatrician, a behavioural technician, a speech pathologist, and a layperson with a background in information technology and a parent. Including the layperson and parent in the group was based on the different benefits outlined by Delbecq and Van de Ven (1971) and Van de Ven and Delbecq (1972). First, it eliminates sole focus on the professional perspective. Secondly, the user's needs and perspective, in this case, the parent's, are included, and finally, it allows for a more robust assessment of the screening tools because of the user's participation and representation in the decision-making process. This professional and public group method was utilised in several health-related studies (McMillan et al., 2015; Tammela, 2013).

Through the first author's networks, experts were either identified through parent networks or recommended by general practitioners who were approached and asked to share information about the study. Seventeen experts, parents, and laypersons were invited via email, telephone messages and personal contact. Participants were given three possible meeting dates and asked to provide feedback on availability. They were followed up via email, telephone calls and chat messages, with several reminders sent to the non-responders. Following telephone and chat responses, the proposed meeting dates and schedules were shared with eight individuals who confirmed their availability. All were provided with the information sheets about the study.

Measures

Autism Spectrum Disorder Screening Tools

The two screening tools reviewed were the Social Communication Questionnaire (SCQ) and the Autism Screening Quotient (AQ-10) adolescent version. Both measures were identified via a systematic review (Nwokolo et al., submitted). The SCQ is a brief 40-item parent or caregiver-report screening measure used widely in research (Berument et al., 1999). The measure has two versions, the lifetime version and the current version, both focusing on symptoms of autism most likely to be observed by the individual's principal caregiver. The AQ-10 is the short version of the AQ-50 (Allison et al., 2012) and is usually completed by a parent or caregiver. The lifetime version of the SCQ and the AQ-10 were presented to the participants.

Intellectual Disability Screening Tools

Two tools identified through the systematic review were presented to the participants (Nwokolo et al., submitted). The tools were the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q) and the Screener for Intelligence and Learning Disabilities (SCIL 14-17). The CAIDS-Q is a

short 7-item screening questionnaire for detecting intellectual disabilities in children and adolescents developed in the UK by McKenzie and Paxton in 2012. The SCIL 14 – 17 was developed as a 14-item screening tool for adults and adolescents in the Dutch language (Nijman et al., 2016; Geijsen et al., 2016). For the original SCIL for adults, a question on the highest level of education must be completed by the respondent. For the SCIL for 14–17-year-olds, this question was changed by the original authors to 'What education (school) are you at now?'.

There is no commercially available English version of the SCIL 14 – 17, so it needed to be translated for this study. Translation from Dutch to English followed the procedure laid out by the International Test Commission (International Test Commission, 2017). To ensure that the overlap in definition and constructs measured were adequately captured, a 2-person expert and bi-lingual team of clinical psychologists in the field of intellectual disability translated the Dutch version to English (both team members were Dutch; one was resident in the United Kingdom and the other in the Netherlands). English-only speaking clinical psychologists reviewed the English version. The English translation was then returned to the Dutch developers to be re-translated back into Dutch. Following the final review by the Dutch developers, the English version was certified for use. Internationally, this back translation and adaptation process is often used to ensure that linguistic equivalence, psychological, and cultural differences are considered (Grisay, 2003; International Test Commission, 2017). Usually, the source version (Dutch) of the text is translated into the intended version (English) and then translated back to the original language for comparison and identification of possible discrepancies. This back-translation technique is useful for detecting essential interpretation issues or mistranslations (Hambleton, 2002; Grisay, 2003). Once the English version correctly reflected the Dutch version's content, structure, and language, the research team finalised the arrangement and utilised it with the Nominal Group.

Procedure

Ethical approval for the study was obtained from the University of Kent, Tizard Centre Ethics Committee, and the National Health Research Ethics Committee of Nigeria (NHREC; NHREC/01/01/2007-16/09/2019).

The consensus meeting started late morning and lasted six hours with a one-hour lunch break. The researcher, who also facilitated, made a 15-minute presentation to provide background information on the project and a summary of the results of the systematic review (Nwokolo et al., submitted) for the participants. Following that, the nominal group process was explained, and the participants were given the consent form to read and sign. Consent included granting permission to record the meeting. The participants were assured that all information would be anonymised and treated confidentially. Signed consent forms indicated a willingness to participate.

Additionally, the researcher explained the goal and expected outcomes to the participants. Once all questions were answered and clarity provided, the screening tools to be reviewed were handed out. The meeting was organised in two sessions: the first segment discussed the autism tools, while the intellectual disability tools were discussed in the second half of the session. As the screening tools were not redesigned, the NGT method was modified (McMillan et al., 2016). Phases one (problem exploration) and two (knowledge exploration) were merged, and the first step – silent generation of ideas - was modified to review each measure's existing format, questions, and content. After that, one measure was selected for autism and one for intellectual disability and reviewed in detail. Phases three to five (priority development, program development and program evaluation) were merged for the second stage. During the second stage, the discussion was open, and group members' interactions were allowed but moderated by the facilitator. Allowing open discussion and interaction was a culture-based decision that had minimal influence on the individual suggestions and conclusions. During the discussions, ideas and comments were stated in a round-robin manner (one participant at a time stated a single idea to the group), with clarifications given. The facilitator collated all suggestions, votes, and agreements. The entire meeting was recorded, transcribed, and analysed for themes.

Measure Selection

The participants received four screening tools (two each for autism spectrum disorder and intellectual disability): the SCQ, AQ-10, CAIDS-Q and SCIL 14 – 17 to review. Participants were asked to assess the face validity, content validity and cultural relevance of all four tools. Following the assessment, a comparison was made between the SCQ and the AQ-10, and the advantages and disadvantages were discussed. Similarly, the group compared the SCIL 14 – 17 to the CAIDS-Q. In-depth discussion of the preferred measures followed with the facilitator's guidance. Ambiguous words and examples were clarified, and more culturally relevant words or phrases were suggested. After the discussion and clarification, the suggested options were voted on and selected.

Data Analysis

Consensus

Although Fink et al. (1984) stated that there are no specific rules for establishing consensus, they describe the various criteria, such as percentage of participants in support, topics with the most votes, and rating on a scale. Fink et al. (1984) also mentioned that the narrower the criteria, the more challenging obtaining consensus usually is. Given that consensus meetings aim to determine the extent of agreement between experts, the threshold for agreement is typically predetermined. Williamson et al. (2012) and Humphrey-Murto et al. (2017) suggested that advance consideration and a clear definition be given to the criteria for consensus. Various thresholds have been reported in the literature as acceptable; 67% (Cantrill et al.,

1996), 75% and 80% (McConachie et al., 2018), while Williamson et al. (2012) suggested 70% for consensus. The extent to which each participant agrees with the contents of each measure under consideration was defined as agreement. Based on Williamson et al. (2012), a criterion of 75% threshold was set for this study. The threshold of 75% meant 6 out of the 8 participants (Fink et al., 1984) had to agree on the retention of the original wording of the measure or with the suggested modification. A simple response tallying for each question was used, and percentage agreement was calculated. For the SCQ, each of the 40 questions was analysed separately and similarly for the 14 questions of the SCIL 14 – 17. All data were collated and analysed using Microsoft Excel for Windows 10.

Meeting Transcription and Theme Generation

Because consensus methods are broadly considered as qualitative methods (Tammela, 2013; Jones & Hunter, 1995), the meeting recording was transcribed and analysed following the thematic analysis (TA) methodology (Braun & Clarke, 2006; Alhojailan, 2012). Thematic analysis has been used before to analyse NGT data (McMillan et al., 2014; Søndergaard et al., 2018). To mitigate against eclipsing individual positions, individual idiosyncrasies are included as themes, with the reverse also being applicable, where the group is not eclipsed while privileging the individual. This study used a combination of the process and modifications outlined in Tomkins & Eatough (2010) and Palmer et al. (2010). Tomkins & Eatough (2010) employed a superordinate (individual level) theme analysis while maintaining the group interactive context. Palmer et al. (2010) explored the participants' experiential claims and concerns, followed by a parallel commentary development in the group discussion context.

The following steps were implemented in analysing the data with an explanation of what was done.

7. Familiarisation with the data. Familiarisation involved the first author transcribing the data and re-reading the transcript at least three times while appraising each participant's comment and contribution. Noting of initial ideas also occurred.
8. Initial codes generated. Codes were generated based on the meaning of each participant's thoughts and were colour coded. Comments were made on the right-hand side of the margin about the meaning.
9. Searching for themes. The colour-coded texts were clustered into potential themes on a group level. Coloured words, phrases, sentences, and passages were re-read to get a sense of the overall perspective from a particular participant without eclipsing the group. Each colour represented an emerging theme.

10. Collating codes into themes. All data were extracted and gathered into relevant main and sub-themes. Main and sub-themes were produced and named (Table 1). These themes are described in some detail with reference to direct quotes from the participants.
11. Reviewing themes. Themes were cross-checked relative to the codes with ongoing analysis to refine the specifics of each theme and the overall story.
12. Producing the report. Examples of effective extracts were selected and analysed for inclusion in the study report. The selection of the extracts was made relative to the research question.

Data trustworthiness is relevant in qualitative research work and Nowell et al. (2017) outlined the process to ensure data trustworthiness. The process expands on the steps outlined in Braun & Clarke (2006). Trustworthiness is measured by credibility, transferability, dependability, and confirmability criteria. In phase 1, for instance, the process requires prolonged engagement with the data and maintaining records of all data field notes and transcripts. The data were reviewed thrice during transcription, with continuous reference to the data while putting together the study report. The raw data and original notes were stored in a secure place. Other steps suggested by Nowell et al. (2017) are team consensus on themes in phase 5, member checking (phase 6) and documentation of meetings (phase 2). The research team members (second & third authors) vetted the themes and sub-themes proposed by the first author and reached an agreement. Additionally, the summary of all meeting sessions was documented and stored via a secure system.

Results

Meeting outcome

Seven out of the eight participants were present at the start of the meeting. The eighth participant joined about 40 minutes later—another participant left due to a prior engagement about an hour before the end of the meeting. Although one participant joined late and another exited early, the agreement calculation was based on the total number of participants, eight. However, this had no significant impact on the results reported in the relevant sections below, as the threshold of 75% agreement set for the study was exceeded (see below for details).

For screening intellectual disability, the participants chose the SCIL 14 – 17 as they found it more robust and thorough, stating that they felt the CAIDS-Q was overly simplified. The group indicated that the SCIL 14 – 17 tested the relevant skills such as intellectual functioning and some adaptive skills.

Similarly, for screening autism spectrum disorder, the SCQ was chosen over the AQ-10 as more robust and comprehensive with questions that examine the relevant autism spectrum domains.

Themes

Following the analysis of the transcript, three themes were identified. Namely language, cultural relevance, and face validity. These are listed in Table 1.

Table 1 About Here

Language

This theme focuses on how Nigerians use the English language and the meaning attached to certain words, sometimes depending on the context. The word '*rituals*' used in question 8 of the SCQ was deemed to have a negative connotation, and the participants advised that an alternative word be used. In the African context and Nigeria, rituals involve sacrifices to 'deities' or some god. The word 'rituals' was therefore changed to 'routines'. Meanwhile, Question 9, 'has her/his facial expression usually seemed appropriate to the particular situation, as far as you could tell?' on the SCQ elicited the following dialogue:

> *R: how do we determine what appropriate facial expression is?*
>
> *BK: to the situation, it says 'to the particular situation'. For instance, someone is dead, and you're smiling.*
>
> *AB: or they're supposed to be afraid or scared*
>
> *AO: again, one of the things I have come to realise is that there is a Nigerian English. If I want to say that thing, I may say that 'has her/his facial expression often reflected the situation at hand', as far you could tell?*

A good number of Nigerian dialects are spoken with a double emphasis, which may appear as either verbal or logical tautology when translated to English. In Yoruba, for instance, the phrase '*pada sẹhin*', when translated to English, means '*return back*'. Thus, AO stated that 'there is a Nigerian English'. Another example was item 14 of the SCQ, 'has she/he ever seemed to be *unusually* interested in the sight, feel, sound, taste, or smell of things or people?'. The discussion was as follows:

AB: sorry to take you back to #14. Even though it cuts across all senses, some persons, when you talk about feeling things, may not be able to relate that to touch. So how do you go about that?'.

Me: the parent or the individual…?

BK: what I hear him say is that the word 'feel' in this context may be interpreted emotionally as opposed to tactile

AB: is there a way to put 'touch' in brackets?

All: tactile

GB: that one is grammar

OO: 'touch' is more appropriate for our environment than tactile

R: 'tactile' sounds really oyibo, 'touch'

AO: there is Nigerian grammar even with academic papers. The editor will ask you to find a native English speaker who will edit, who knows exactly what you are saying but puts it in a different way. But when you are dealing with instruments like this, I believe the more you 'Nigerianise' it, the more you'll get the appropriate response

Further discussion considered Question 31 of the SCQ, which asks, 'when she/he was 4 to 5, did she/he ever try to comfort you if you were sad or hurt?'. Since Nigerians say 'sorry' for nearly every incident, including those the individual is not responsible for, the group recommended adding examples for clarity.

For the SCIL 14 – 17, language reference was minimal. The agreement was to change the word 'GP' to 'doctor' as the term 'GP' is not used in Nigeria. Regarding the dictation component of the measure in question 12, the group agreed to exclude words with consonants likely to be mispronounced to avoid possible h-dropping (such as hitting). Question 13 of the SCIL 14 – 17 tests reading skills, and the ability to read fluently incorporates the reader's comprehension, familiarity with the words and background knowledge of the context. Question 13 in the SCIL 14 – 17 includes these words: '*pay for parking by mobile phone. When you have parked your car, log in on your mobile using the (location) code as advertised/displayed on the signs and parking machines. When you leave, you log out by phone/mobile*.' In the Nigerian environment, parking is not paid for like this. However, an approximate equivalent is the point of sale (POS) machines with bank cards used in stores. To use language that will be familiar in the Nigerian context, the group agreed on 'bank card'. Below are some excerpts from the discussion.

BK: this is based on those places where you have parking metres. Then you slot in and pay for your parking. Where there's no context for it…

AO: parking at the mall. No, you don't even need to do the possibility. Just say, the process of…

R: are you allowed to change it completely?

BK: no, you're turning it into a title. It's a sentence. It tells you it is possible to pay by phone, then it now telling you how to do it.

Me: (tell my story). In this context, in order not to change the story completely, we can say "ATM" or "POS".

R: can we say 'card'? Is it everybody that knows 'POS'?

BK: yes, is it not every Nigerian that knows "POS"?

R: adolescents?

BK: yeah, it's the language of the environment.

AO: yes, it is. "POS" is the language, but I don't want us to introduce a word that is not actually a word; "POS".

Me: ok, so, with 'card' because 'POS' is 'point of sale'. So, by 'card'.

Cultural Relevance

There were three sub-themes under cultural relevance: the examples given, family dynamics (the way parents related with their children), and context. All the examples given related predominantly to the design environment, the West. The group advised utilising more culturally relevant examples. For instance, vacuuming, gardening, or mending things were given as examples in question 21 on the SCQ. In the Nigerian context, not everyone vacuums, and mending things appeared vague. Therefore, the participants suggested using examples such as sweeping and washing. A portion of the dialogue follows below.

AO: sweeping, more people sweep than they vacuum even if they are cosmopolitan or whatever group we are looking at

BK: maybe just cleaning, washing

OO: that's appropriate. Just look at things that we do here

GB: local content

Following the discussion on question 21, the group agreed that the questions were relevant and appropriate from questions 32 and below. However, for some questions, more local examples and songs were suggested as replacements of Western ones. GB mentioned activities such as 'backing a baby' (a traditional African method where mothers carry babies and infants on their backs swathed in cloth), 'cooking with hibiscus flower', 'playing mummy and daddy'. At the same time, AO said, *"I see that even in real practice, what differentiates what we do at times from questionnaires alone, is that opportunity to spend time explaining what we do, unlike just giving it to them to fill. You realise that the more you are engaging, the more the individual is able to know exactly what you are talking about."* Buttressing AO's point, AB said, *"which is what I've found with parents most often. When you give them a questionnaire*

like this, what they do is to fill, and when they get to where they don't understand, they will ask a question. Once you give them examples, it's clear, and they give you other examples." AO, *"so meaning that a useful questionnaire in this environment will do well to have short-short examples where necessary, which is what we are doing."*

Turning to family dynamics, question 2 on the SCQ designed to assess the extent of vocalisation may require some explaining as holding 'to and fro conversation' is not the norm in typical Nigerian homes. Although there is some shift regarding this, children are often expected to respond to questions asked by parents, rather than engage in 'chit chat'. The younger parents are at the fore of changing this narrative. One of the younger participants, AO, said, *"to and fro, they may get a little bit but once you say converse (pause), in fact, a lot of people complain that they are coming to come and tell you that yeah, they are talking, but he is still having problems with conversation."* Therefore, the group agreed to leave the question as is and give examples of what a 'to and fro conversation' entails.

Face Validity

The last theme, face validity, covers environment and professional versus parent's understanding of a question. As the SCQ is a self-administered (parent) questionnaire, the participants opined that it might be more useful if the professionals administered it to allow for explanations where there is the possibility of confusion or lack of clarity. For instance, question 4 reads as *'has she/he ever used socially inappropriate questions or statements? For example, has she/he ever regularly asked personal questions or made personal comments at awkward times?'* To which the following dialogue ensued.

> *GB: when a child is done eating, there is no need to say, 'will the food be ready'.*
>
> *AO: I'm thinking that while I agree that it is clear, we must also remember that if you are very familiar with ASD, some of these questions will be clear to you. But if you are not familiar with ASD, you may not actually grasp it. This particular question, we all know what this question is trying to test.*
>
> *Me: that's why I'm looking at my parent; as a parent, if you are given this question, is this clear enough. Are you able to answer yes, or no?*
>
> *R: yes, but I am a parent who already knows quite a bit. Going back to what he's saying, I am not a lay parent that has just come.*
>
> *BK: the idea of inappropriateness, from the example given, it's one more out of context versus something more socially inappropriate in terms of asking a personal question.*

Once all participants had expressed their opinions, the group agreed that the correct response, 'yes' or 'no', would be elicited from respondents irrespective of their background.

Regarding the screening tool for intellectual disability, the SCIL 14 – 17, the group discussed questions 1 and 2 extensively. The questions are centred on special education and level of education. Many Nigerian schools in the urban areas purportedly offer special education services.

> *OO: looking at question 1 for me, looks like the first stem and second stem are looking at the same thing.*
>
> *GB: but in the true sense of it for people practising, for example, 'did you receive special education?' You can be in a regular school system and be receiving support from a unit.*
> *PA: yes, and you're receiving support from a unit. Yes.*
> *GB: do you go to a special needs school? You could have a school that is a special needs school, all the teachers there are specialist trained personnel, and you have special materials, and that school is labelled for that specific learning difficulty. It may be school for hearing impaired, school for individuals with learning disability or school for individuals with autism. So, you could have that, or did you have a special education need? That means are you having challenges with learning, typically. So, the three questions are not actually the same. We could sample different people differently.*
> *AO: in any case, the answer is 'yes' or 'no'. Meaning that when you read through the question, anyone of it is what you are responding to.*
> *GB: you will fall into one category. The one that applies to you.*

Once the different educational categories and services were agreed on, the team accepted the questions. The levels of education were also expanded to include the different curricula, both national and international, offered in the country. Some of these are the West African Examination Council (WAEC), polytechnic, monothecnic and teachers' colleges.

In discussing question 3, *'do you receive or have you received support from a service for people with Intellectual Disability (excluding a home tutor or lesson teacher)*?' was examined at length by the participants. The exclusion of home tutors and lesson teachers was the consensus as there are no such services for a person with intellectual disability in Nigeria. A private tutor (lesson teacher) is typically employed once a child struggles in school. However, some who do not struggle with schoolwork have these tutors as a competitive advantage. The distinction lies in their academic performances, so having a private tutor does not necessarily indicate a pupil is struggling to understand the material.

For both the SCQ and the SCIL 14 – 17, the consensus from the group was that face validity was met. The items on the SCIL 14 – 17 have specific and relevant questions that test for intellectual disability. At

the same time, the different DSM-5 domains (social and communication deficits, repetitive and restricted behaviours) of autism spectrum disorder are captured in the SCQ.

Social Communication Questionnaire (SCQ) for screening Autism Spectrum Disorder

The participants agreed that several changes were needed, such as that more local and culturally relevant examples should be given in the tool. For instance, for question 6 ("has she/he ever used words that she/he seemed to have invented or made up her/himself; put things in odd, indirect ways; or used metaphorical ways of saying things (e.g., saying *hot rain* for *steam*)?"), experts' opinion was to give examples to the respondents in context with Nigeria. Thus, for question 6, the replacement for "*hot rain for steam*" would be "*jagbajantis for mess*".

Another example was question 8; the word 'ritual' was explained as 'routine' to remove any fetish connotation. According to Hambleton (1996, p28), "when an instrument is adapted for use in another population, documentation of the changes should be provided, along with evidence of the equivalence." The list of examples of other culturally relevant words, examples, and clarifications are in Table 2. Overall, the participants agreed that 23 (58%) of the 40 items were culturally relevant and required no modification. After discussions and adaptations, between 87.5% and 100% agreement were achieved for all 40 questions. Table 3 shows the SCQ questions which were modified.

_____

Table 2 About Here

_____

_____

Table 3 About Here

_____

Screener for Intelligence and Learning Disabilities (SCIL 14 – 17) for screening Intellectual Disability

The same participants reviewed the autism spectrum disorder and intellectual disability measures. The agreement for the SCIL 14 – 17 to give more contextual and culturally relevant examples was between

87.5% and 100%. Thus, more contextual and culturally relevant examples were given, in addition to including other relevant educational categories. There is no commercially available English version of the SCIL 14 – 17, and this study was an effort to create one. Therefore, in examining the face validity, culturally relevant words and examples were included. Question 1 on the level of education was modified to include all the different categories of educational qualifications obtained in Nigeria. One key factor was language. In most western societies, a 'diploma' refers to a secondary school certificate, while in Nigeria, a 'diploma' refers to certificates obtained in post-secondary school. In question 3, because there are no 'services' as obtained in the West, 'services' had to be modified to exclude individuals who provided extra tutoring at home as a competitive advantage. However, where individuals visited any psychiatric facility or psychologist, these qualified as receiving service. Another example is changing the word 'GP' to 'doctor' as the term 'GP' is not utilised in Nigeria. Results of other modifications are provided in Table 4. Table 5 shows the old and modified questions for the SCIL 14 – 17.

Table 4 About Here

Table 5 About Here

Discussion

Cross-cultural adaptation of any tool is often complicated, thus requiring careful elimination of possible construct, item, and method biases (Van de Vijver & Poortinga, 1997; Van de Vijver & Tanzer, 2004). Beyond the biases identified by Van de Vijver & Poortinga (1997) and Van de Vijver & Tanzer (2004), Peña (2007) identified another type of bias which can occur when conducting cross-cultural adaptation of screening tools called 'equivalence'. According to Peña (2007), there are four types: cultural, linguistic, metric and functional equivalence. A qualitative review of the dialogue between the nominal group participants revealed that the biases of concern were around linguistic, cultural, and functional

equivalence. The linguistic equivalence ensures the consistency of words, sentences, meaning, and language used between the original and the adapted tool (Peña, 2007). One challenge with linguistic equivalence is that even when words are the same across the original and adapted tools, culture, interpretation, and word familiarity may result in potential differences in patterns of responses. In the SCIL 14 – 17, for instance, the phrase 'mobile phones' is similar in the Nigerian context; however, the function attributed to it was different. With cultural equivalence, how members of different linguistic and cultural groups interpret the underlying meaning of words or items is crucial. For instance, question 2 in the SCQ asks about 'holding to and fro conversation', which is not the norm in an average Nigerian family.

With functional equivalence, both the original tool and the adapted version should allow examination of the same construct. Both versions should offer the same opportunity to demonstrate knowledge while eliciting the intended response from participants. An example of this was the observation made on question 14 of the SCQ on 'tactile' in the original version versus 'touch' in the Nigerian context. Overall, there is an interaction between the linguistic, cultural, and functional equivalence which should not be ignored in the adaptation process. Additionally, participants were concerned about the method bias (mode of administration) and item bias, especially for the SCQ. In cultures where social interactions and dialogues are salient, dyadic administrations may be more valuable.

The modified nominal group technique was used to select the most robust screening tool for autism spectrum disorder and/or intellectual disability from those identified through a systematic review (Nwokolo et al., submitted). The cultural relevance, face validity, and content validity for use with the Nigerian adolescent were examined. The Social Communication Questionnaire is an existing measure developed in the Western environment with various translations. Three participants were familiar with the SCQ and used it often. The group reviewed the Lifetime English version with consensus reached on all the face and content validity items. On cultural relevance, the consensus was to use indigenous examples in language and activities mentioned in the SCQ. The group agreed that although the SCQ is a self-administered tool, it may be best administered as a quasi-interview questionnaire to get a more accurate response in the Nigerian context. Doing so will allow the administrator to explain potentially confusing concepts, quickly substitute examples, expound phrases, and note areas of importance or value to the respondent. This view of adapting tools to meet the specific culture and environment of intended use was captured by Soto et al. (2015).

Reviewing the Screener for Intelligence and Learning Disabilities required more depth as there currently is no English version. The group chose the SCIL 14 – 17 over the CAIDS-Q, stating that the SCIL 14 –

17 had specific questions in certain areas like mathematics and reading. The SCIL 14 – 17 was deemed more engaging and functional. Not only did they agree on the face validity, but the group also noted that the contents of the SCIL 14 – 17 tested individual abilities and the DSM-5 domains for ID (conceptual, social, and practical). The question on the level of education was expanded to include all the different curricula offered in Nigeria, including the Nigerian, British and American curricula.

To have an adapted tool that is culturally relevant, linguistically appropriate, and applicable to the environment of intended use, such as Nigeria, individuals who understand the people and are also familiar with the construct of interest need to be involved in the adaptation process. Whereas adaptation of tools includes language translation, modification of methods, clarification of concepts, and sometimes changing the content, for the tool to be genuinely relevant culturally, the values and peculiarities of the environment of intended use should be considered (Al Maskari et al., 2018; Soto et al., 2015). For instance, the word '*ritual*' in the SCQ will elicit a different response as some people believe in idols and engage in 'rituals' (sacrificial killings) in Nigeria. Therefore, respondents will likely answer 'no' if they do not hold such beliefs or ignore the question where they feel it is a private event. The nominal group paid attention to such content and recommended that alternative wording be used or have the administrator explain the question. Similarly, on the SCIL 14 – 17, the group suggested that the type of 'service' be qualified. In Nigeria, there are no similar government-funded organisations or services like those in the West, where the tool was initially developed.

With the increasing global awareness of developmental disabilities such as ASD and ID (Malcolm-Smith et al., 2013), more individuals, especially younger children, now have early screening and intervention. Literature, however, remains extant on research involving older children and adolescents. The narrative on younger children is no different in Africa, especially in countries like Nigeria. However, the significant difference and challenges are with the older children and adolescents who have had no access to screening either by design or parents' choice. The lack of early identification leads to poor social integration, reduced quality of life and lack of intervention (Bargiela et al., 2016; Nwokolo et al., submitted). Addressing the adolescent screening gap requires robust and culturally relevant measures with face and content validity. Resources in terms of financing and expertise are also potential barriers to developing new screening tools for low to middle-income economies; thus, adapting an existing tool is a prudent option. Substantial research on the adaptation of screening tools has been conducted in the West and other medium-income economies, where it is recognised that cultural disparities potentially impact adaptation (Long et al., 2020; Grinker et al., 2015). However, very little work has been done in Africa and other low to middle-income economies.

Limitations

This study focused on adapting screening tools in English, which is spoken widely in Nigeria. However, even in Nigeria, this may make the tools difficult to use in rural areas where the local dialects such as Igbo, Hausa, Yoruba and Ijaw are spoken. Moreover, the tools would not necessarily be helpful in non-English speaking African countries, and they would, in any case, need further cultural adaptation for use elsewhere.

While every effort was made to involve a range of professionals with relevant experience, as well as family members and a layperson in the consensus group, the late arrival of one of the participants and the early exit of another meant that expert representation in those fields was not available for the entire meeting period.

While efforts to ensure the qualitative data's trustworthiness were embarked on, it is possible that data validity was not explored in its entirety. One possible means of exploring data validity would have been the use of other investigators (different colleagues in the same field analyse the data using the same qualitative method or focus groups) or theory triangulation (use of multiple perspectives in interpreting the data) (Guion et al., 2011).

Finally, using an evaluation team outside the researcher's group may have lent different perspectives or interpreted the data differently.

Conclusion

Realising that the adaptation process beyond language translation can be complicated and challenging, using the appropriate knowledge, skill, and expertise is crucial. A group of Nigerian experts in the relevant professions were consulted to review the four identified tools for screening for ID and autism, for face validity, content validity and cultural validity, with two tools chosen for scrutiny and adaptation. Assessing some of the properties (face validity, content validity) of the screening tools using the NGT was useful. Following the recommendations and consensus of the group, the SCQ and the SCIL 14 – 17 were agreed on as measures to validate with the Nigerian adolescents, with only a small number of adjustments to allow for different use of language, customs and environment in the Nigerian context.

Acknowledgements

The authors would like to thank Xavier Moonen and Anne-Marie for the translation work on the SCIL.

Financial Disclosure

None

Funding Source

This research received no specific grant from the public, commercial, or not-for-profit funding agencies.

Conflict of Interest

The authors do not have any conflict of interest to declare, apart from Xavier Moonen, who is one of the developers of the Dutch version of the SCIL. He receives an annual financial contribution from the publisher for, among other things, further research.

References

Al Maskari, T. S., Melville, C. A., & Willis, D. S. (2018). Systematic review: Cultural adaptation and feasibility of screening for autism in non-English speaking countries. *International Journal of Mental Health Systems, 12*(1), 1-19.

Alhojailan, M. I. (2012). Thematic analysis: A critical review of its process and evaluation. *West East Journal of Social Sciences, 1*(1), 39-47.

Bargiela, S., Steward, R., & Mandy, W. (2016). The experiences of late-diagnosed women with autism spectrum conditions: An investigation of the female autism phenotype. *Journal of Autism and Developmental Disorders, 46*(10), 3281-3294.

Bartunek, J. M., & Murninghan, J. K. (1984). The Nominal Group Technique: Expanding the basic procedure and underlying assumptions. *Group & Organization Studies, 9*(3), 417-432.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186-3191.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism Screening Questionnaire: Diagnostic validity. *British Journal of Psychiatry, 175*(5), 444-451.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Canal-Bedia, R., García-Primo, P., Martín-Cilleros, M. V., Santos-Borbujo, J., Guisuraga-Fernández, Z., Herráez-García, L., del Mar Herra´ez-Garcı´a, M., Boada-Muñoz, L., Fuentes-Biggi., J & Posada-de La Paz, M. (2011). Modified Checklist for Autism in Toddlers: Cross-cultural adaptation and validation in Spain. *Journal of Autism and Developmental Disorders, 41*(10), 1342-1351.

Cantrill, J., Sibbald, B., & Buetow, S. (1996). The Delphi and Nominal Group Techniques in health services research. *International Journal of Pharmacy Practice, 4*(2), 67-74.

Cuesta-Gómez, J. L., Andrea Manzone, L., & Posada-De-La-Paz, M. (2016). Modified Checklist for Autism in Toddlers cross-cultural adaptation for Argentina. *International Journal of Developmental Disabilities, 62*(2), 117-123.

Delbecq, A. L., & Van de Ven, Andrew H. (1971). A group process model for problem identification and program planning. *The Journal of Applied Behavioral Science, 7*(4), 466-492.

Delbecq, A. L., Van de Ven, Andrew H, & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to Nominal Group and Delphi processes*. Glenview, IL: Scott, Foresman and Company.

Delbecq, A. L. (1967). The management of decision-making within the firm: Three strategies for three types of decision-making. *Academy of Management Journal, 10*(4), 329-339.

Fink, A., Kosecoff, J., Chassin, M., & Brook, R. H. (1984). Consensus methods: Characteristics and guidelines for use. *American Journal of Public Health, 74*(9), 979-983.

García-Primo, P., Hellendoorn, A., Charman, T., Roeyers, H., Dereu, M., Roge, B., Baduel, S., Muratori, F., Narzisi, A., Van Daalen, E., Moilanen, I., Posada-de la Paz, M., & Canal-Bedia, R. (2014).

Screening for autism spectrum disorders: State of the art in Europe. *European Child & Adolescent Psychiatry, 23*(11), 1005-1021.

Geijsen, K., Kop, N., & de Ruiter, C. (2018). Screening for intellectual disability in Dutch police suspects. *Journal of Investigative Psychology and Offender Profiling, 15*(2), 200-214.

Grinker, R. R., Kang-Yi, C. D., Ahmann, C., Beidas, R. S., Lagman, A., & Mandell, D. S. (2015). Cultural adaptation and translation of outreach materials on autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(8), 2329-2336.

Grisay, A. (2003). Translation procedures in OECD/PISA 2000 international assessment. *Language Testing, 20*(2), 225-240.

Guion, L. A., Diehl, D. C., & McDonald, D. (2011). Triangulation: Establishing the validity of qualitative studies. *Edis, 2011*(8), 3-3.

Hambleton, R. K. (1996). Guidelines for adapting educational and psychological tests. *Paper Presented at the Annual Meeting of the National Council on Measurement in Education (New York, NY, April 9-11, 1996).* Retrieved from https://files.eric.ed.gov/fulltext/ED399291.pdf

Hambleton, R. K. (2002). Adapting achievement tests into multiple languages for international assessments. In Porter, AC and Gamoran, A. (Editors) Methodological Advances in Cross National Surveys of Educational Achievement. (pp. 58-79). Washington DC: National Academy Press Washington.

Humphrey-Murto, S., Varpio, L., Gonsalves, C., & Wood, T. J. (2017). Using consensus group methods such as Delphi and Nominal Group in medical education research. *Medical Teacher, 39*(1), 14-19.

Hutchings, H. A., Rapport, F. L., Wright, S., Doel, M. A., & Wainwright, P. (2010). Obtaining consensus regarding patient-centred professionalism in community pharmacy: Nominal Group work activity with professionals, stakeholders and members of the public. *International Journal of Pharmacy Practice, 18*(3), 149-158.

Hutchings, H., Rapport, F., Wright, S., Doel, M., & Jones, A. (2012). Obtaining consensus about patient-centred professionalism in community nursing: Nominal Group work activity with professionals and the public. *Journal of Advanced Nursing, 68*(11), 2429-2442.

International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests (Second Edition)*, Retrieved from www.InTestCom.org

Jones, J., & Hunter, D. (1995). Consensus methods for medical and health services research. *BMJ (Clinical Research Ed.), 311*(7001), 376-380.

Long, K. A., Gordillo, M., & Orsmond, G. I. (2020). Improving the validity and generalizability of adult autism research through incorporating family and cultural contexts. *Autism in Adulthood, 2*(3), 177-184.

Malcolm-Smith, S., Hoogenhout, M., Ing, N., Thomas, K. G., & de Vries, P. (2013). Autism spectrum disorders—Global challenges and local opportunities. *Journal of Child & Adolescent Mental Health, 25*(1), 1-5.

Matsumoto, D., & Yoo, S. H. (2006). Toward a new generation of cross-cultural research. *Perspectives on Psychological Science, 1*(3), 234-250.

McConachie, H., Mason, D., Parr, J. R., Garland, D., Wilson, C., & Rodgers, J. (2018). Enhancing the validity of a quality of life measure for autistic people. *Journal of Autism and Developmental Disorders, 48*(5), 1596-1611.

McKenzie, K., Paxton, D., Murray, G., Milanesi, P., & Murray, A. L. (2012). The evaluation of a screening tool for children with an intellectual disability: The Child and Adolescent Intellectual Disability Screening Questionnaire. *Research in Developmental Disabilities, 33*(4), 1068-1075.

McKenzie, K. and Paxton D. (2012). *Child and Adolescent Intellectual Disability Screening Questionnaire*. GCM Records, Edinburgh.

McMillan, S. S., Kelly, F., Sav, A., Kendall, E., King, M. A., Whitty, J. A., & Wheeler, A. J. (2014). Using the nominal group technique: How to analyse across multiple groups. *Health Services and Outcomes Research Methodology, 14*(3), 92-108.

McMillan, S. S., Kelly, F., Sav, A., Kendall, E., King, M. A., Whitty, J. A., & Wheeler, A. J. (2015). Consumers and carers versus pharmacy staff: Do their priorities for Australian pharmacy services align? *The Patient-Patient-Centered Outcomes Research, 8*(5), 411-422.

McMillan, S. S., King, M., & Tully, M. P. (2016). How to use the Nominal Group and Delphi techniques. *International Journal of Clinical Pharmacy, 38*(3), 655-662.

Mokkink, L. B., De Vet, H. C., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN risk of Bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research,* 27(5), 1171–1179.

Murphy, M., Black, N., Lamping, D., McKee, C., Sanderson, C., Askham, J., & Marteau, T. (1998). Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment (Winchester, England), 2*(3), i-88.

Nah, Y., Young, R. L., Brewer, N., & Berlingeri, G. (2014). Autism Detection in Early Childhood (ADEC): Reliability and validity data for a level 2 screening tool for autistic disorder. *Psychological Assessment, 26*(1), 215.

Nijman, H., Kaal, H., van Scheppingen, L., & Moonen, X. (2018). Development and testing of a Screener for Intelligence and Learning Disabilities (SCIL 14-17)). *Journal of Applied Research in Intellectual Disabilities, 31*(1), e59-e67.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1609406917733847.

Nwokolo, E. U., Langdon, P. E., & Murphy, G. H. (submitted). Screening for intellectual disabilities and/or autism amongst older children and young adults: A systematic review of tools for use in Africa.

O'Neil, M. J., & Jackson, L. (1983). Nominal Group Technique: A process for initiating curriculum development in higher education. *Studies in Higher Education, 8*(2), 129-138.

Palmer, M., Larkin, M., de Visser, R., & Fadden, G. (2010). Developing an Interpretative Phenomenological Approach to focus group data. *Qualitative Research in Psychology, 7*(2), 99-121.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development, 78*(4), 1255-1264.

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research,* 27(5), 1147–1157.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131-144.

Søndergaard, E., Ertmann, R. K., Reventlow, S., & Lykke, K. (2018). Using a modified nominal group technique to develop general practice. *BMC Family Practice, 19*(1), 1-9.

Soto, S., Linas, K., Jacobstein, D., Biel, M., Migdal, T., & Anthony, B. J. (2015). A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism,* 19(6), 646-661.

Tammela, O. (2013). Applications of consensus methods in the improvement of care of paediatric patients: A step forward from a 'good guess'. *Acta Paediatrica, 102*(2), 111-115.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018a). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research*, 27(5), 1159–1170.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., De Vet, H., Bouter, L. M., Alonso, J., Westerman. M. J., Patrick, D. L., & Mokkink, L. B. (2018b). COSMIN methodology for assessing the content validity of PROMs–user manual. *Amsterdam: VU University Medical Center,* Netherlands.

Tomkins, L., & Eatough, V. (2010). Reflecting on the use of IPA with focus groups: Pitfalls and potentials. *Qualitative Research in Psychology, 7*(3), 244-262.

Van de Ven, A. H., & Delbecq, A. L. (1972). The Nominal Group as a research instrument for exploratory health studies. *American Journal of Public Health, 62*(3), 337-342.

Van de Vijver, Fons J. R., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment, 13*(1), 29-37.

Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. European Review of Applied Psychology, 54(2), 119-135.

Williamson, P. R., Altman, D. G., Blazeby, J. M., Clarke, M., Devane, D., Gargon, E., & Tugwell, P. (2012). Developing core outcome sets for clinical trials: Issues to consider. *Trials, 13*(1), 1-8.

Young, R. (2007). *Autism Detection in Early Childhood (ADEC) manual*. Camberwell: ACER Press.

Table 1

Main themes and sub-themes for the SCQ and SCIL 14-17)

Main Theme	**Sub-themes**
Language	• Use of words • Meaning of the word • Context • Nigerian parlance
Cultural relevance	• Examples given • Family dynamics (the way parents relate with their children) • Context
Face Validity	• A professional versus the parent's understanding of the question • Environment

Table 2

List of questions and the agreed cultural examples and modifications for the SCQ

SCQ Item number	Number of votes	Comments and suggested clarifications
1	7	Include examples such that it is clearer (mummy see, etcetera. # of words).
6	8	Give examples of respondents in context Nigeria, e.g., 'jagbajantis' for mess
8	8	For the word 'rituals' use "routine"
9	8	Example laughing when something is funny or showing concern when something is wrong
12	8	Include a 2nd example - combing the doll's hair over and over, switching a torch on and off
13	8	Examples are male dominant; add dolls, etcetera. for females
14	8	For feel put 'touch' in brackets
15	8	Include 'face'
16	8	Examples - hanging upside from a chair, twisting their body into a funny shape, any unusual body movement
18	8	Give other examples - cars, dolls, something that seems like a favourite item
20	8	The words in bracket meant for clarification ('rather than to get something'), we can use "only to get something"
21	8	Local examples such as sweeping, cleaning the table, washing plates
28	8	For engage, put "get" & "keep" in brackets
30	8	Add, e.g., playing hide and seek
31	8	Add e.g., "say sorry"
33	8	Example in brackets (sad, etc.)
34	8	Examples of local songs and common ones; "if you're happy", "ABCD…", “twinkle twinkle", "xxx is a good girl or boy"
35	8	Example playing daddy & mummy, backing a baby*

**a traditional African method where mothers carry babies and infants on their backs swathed in cloth*

Table 3

List of old questions and their modifications for the SCQ

SCQ Item number	**Old question**	**Modified question**
1	Is she/he able to talk using short phrases or sentences? If *no,* skip to question 8.	Is she/he able to talk using short phrases or sentences? If *no,* skip to question 8. How many words can she/he use when talking? For example, 'mummy see', 'come here', 'what is your name?'
6	Has she/he ever used words that she/he seemed to have invented or made up her/himself; put things in odd, indirect ways; or used metaphorical ways of saying things (e.g., saying *hot rain* for steam)?	Has she/he ever used words that she/he seemed to have invented or made up her/himself; put things in odd, indirect ways; or used metaphorical ways of saying things (e.g., saying *jagbajantis* for mess)?
8	Has she/he ever had things that she/he seemed to have to do in a very particular way or order or rituals that she/he insisted that you go through?	Has she/he ever had things that she/he seemed to have to do in a very particular way or order or routines that she/he insisted that you go through?
9	Has her/his facial expression usually seemed appropriate to the particular situation, as far as you can tell?	Has her/his facial expression usually seemed appropriate to the particular situation, as far as you can tell? For example, laughing when something is funny or showing concern when something is wrong.
12	Has she/he ever seemed to be more interested in parts of a toy or an object (e.g., spinning the wheels of a car), rather than using the object as it was intended?	Has she/he ever seemed to be more interested in parts of a toy or an object (e.g., spinning the wheels of a car, combing the doll's hair over and over, switching a torch on and off), rather than using the object as it was intended?
13	Has she/he ever had any special interests that were unusual in their intensity but otherwise appropriate for her/his age and peer group (e.g., trains, dinosaurs)?	Has she/he ever had any special interests that were unusual in their intensity but otherwise appropriate for her/his age and peer group (e.g., trains, dinosaurs, dolls, clothes)?
14	Has she/he ever seemed to be unusually interested in the sight, feel, sound, taste, or smell of things or people?	Has she/he ever seemed to be unusually interested in the sight, feel (touch), sound, taste, or smell of things or people?
15	Has she/he ever had any mannerisms or odd ways of moving her/his hands or fingers, such as flapping or moving her/his fingers in front of her/his eyes?	Has she/he ever had any mannerisms or odd ways of moving her/his hands or fingers, such as flapping or moving her/his fingers in front of her/his eyes or face?
16	Has she/he ever had any complicated movements of her/his whole body, such as	Has she/he ever had any complicated movements of her/his whole body, such as spinning, repeatedly bouncing up and down, hanging upside from a chair, twisting their

	spinning or repeatedly bouncing up and down?	body into a funny shape, any unusual body movement?
18	Has she/he ever had any objects (other than a soft toy or comfort blanket) that she/he had to carry around?	Has she/he ever had any objects (other than a cars, dolls, something that seems like a favourite item) that she/he had to carry around?
20	When she/he was 4 to 5, did she/he ever talk with you just to be friendly (rather than to get something)?	When she/he was 4 to 5, did she/he ever talk with you just to be friendly (rather than only to get something)?
21	When she/he was 4 to 5, did she/he ever *spontaneously* copy you (or other people) or what you were doing (such as vacuuming, gardening, or mending things)?	When she/he was 4 to 5, did she/he ever *spontaneously* copy you (or other people) or what you were doing (such as sweeping, cleaning the table, washing plates)?
28	When she/he was 4 to 5, did she/he ever show you things that interested her/him to engage your attention?	When she/he was 4 to 5, did she/he ever show you things that interested her/him to engage (get & keep) your attention?
30	When she/he was 4 to 5, did she/he ever seem to want you to join in her/his enjoyment of something?	When she/he was 4 to 5, did she/he ever seem to want you to join in her/his enjoyment of something (e.g., playing hide and seek)?
31	When she/he was 4 to 5, did she/he ever try to comfort you if you were sad or hurt?	When she/he was 4 to 5, did she/he ever try to comfort you if you were sad or hurt (e.g., say sorry)?
33	When she/he was 4 to 5, did she/he show normal range of facial expressions?	When she/he was 4 to 5, did she/he show normal range of facial expressions (e.g., sad, angry, happy etc.)?
34	When she/he was 4 to 5, did she/he ever spontaneously join in and try to copy the actions in social games, such as *The Mulberry Bush* or *London Bridge Is Falling Down*?	When she/he was 4 to 5, did she/he ever spontaneously join in and try to copy the actions in social games, such as *ABCD, Twinkle Twinkle Little Star, If You're Happy and You Know It Clap Your Hands*, or *XXX is a good girl or boy*?
35	When she/he was 4 to 5, did she/he play any pretend or make-believe games?	When she/he was 4 to 5, did she/he play any pretend or make-believe games (e.g., playing daddy & mummy, backing a baby)?*

**a traditional African method where mothers carry babies and infants on their backs swathed in cloth*

Table 4

List of questions and the agreed cultural examples and modifications for the SCIL 14 – 17

SCIL 14 – 17) item number	**Number of votes**	**Comments and suggested clarifications**
1	8	Type SEN in full
2	8	Add WAEC/*I*GSE/ SAT, college / monotechenic/ polytechnic/ university
3	8	Write "ID" in full; for - "service" (exclude lesson teachers)
4	8	"In case of emergency or difficult situation…"
5	7	Add Naira sign, change 6,95 to 6.50
6	7	Change GP to Doctor (can use a different context)
7	7	Change GP to Doctor (can use a different context)
8	7	Remove "say every letter"
9	7	"paper" be more specific (newspaper)
10	7	Change to "raining cats & dogs", "make hay while the sun shines", "a stitch in time saves nine"
11	7	Put in boxes
12	7	Change “deer” to “cow”, use "avoid", change "hitting" to "knocking down"
13	7	Change mobile phone to "card"
14	7	Add "mins" to 15, use "detailed"

Table 5

List of old questions and their modifications for the SCIL 14 – 17)

SCIL 14 – 17) item number	**Old question**	**Modified question**
1	Did you receive special education? Do you go to a special needs school? Did you have a SEN?	Do you receive special education? Do you go to a special needs school? Do you have a special educational need (SEN)?
2	Which school/college do you attend now? None Primary school Special needs school GCSE A Level Polytechnic college University Other	Which school/college do you attend now, or did you attend in the past? None Primary school Special needs school WAEC/IGCSE/SAT A-level Polytechnic/Monotechnic/Teacher's college University Other
3	Have you received support from a service for people with ID?	Do you receive or have you received support from a service for people with Intellectual Disability (excluding a home tutor or lesson teacher)?
4	Have you got family members or relatives who you can contact if you have a problem?	Have you got family members, relatives or friends who you can contact if you have a problem (for example a difficult situation or emergency)?
6	Imagine you are at your GP (General Practitioner) 19th of January. He wants to see you again in three weeks. When (which date) would that be?	Imagine you are at your Doctors on the 19th of January. He or she wants to see you again in three weeks. When (which date) would that be?

7	Imagine you are at your GP (General Practitioner) January 3rd. He wants to see you again in three weeks. When (which date) would that be?	Imagine you are at your Doctors on the 3rd of January. He or she wants to see you again in three weeks. When (which date) would that be?
9	Do you read a paper or magazine? If so, which one?	Do you read a newspaper or magazine? If so, which one?
10	What does this mean: The apple doesn't fall far from the tree?	What does this mean: "Like father, like son?"
12	I'm going to read a few sentences for you to write in the box. Try to do this well/correct and as fast as you can. a) We are dumping the load of soil/sand at the back of our house. b) During the night the driver had to swerve/avoid hitting a deer with big antlers.	I'm going to read a few sentences for you to write in the box. Try to do this well or correct and as fast as you can. a. We are dumping a load of sand in the back garden. b. During the night the driver had to avoid knocking down a cow.
13	I'm going to ask you to read a story. Read this as quickly as you can without making mistakes. It is possible to pay for parking by text(phone). When you have parked your car, log in on your mobile using the (location) code as advertised on the signs and parking machines. When you leave you log out by phone.	I'm going to ask you to read a story. Read this as quickly as you can without making mistakes. It is possible to pay for parking with your bank card. When you have parked your car, you use your bank card to pay as advertised/displayed on the signs and parking machines. When you leave you take your receipt.
14	In this box draw a clock that says 9.45 (15 to ten). Draw this as complete/detailed as you can with hands	In this box draw a clock that says 9:45 (quarter to ten). Draw this as complete/detailed as you can with hands.

Appendix 22 – Validation of the SCIL – Submitted Manuscript

A version of this paper was submitted to the Journal of Intellectual Disability Research, in March 2023, as Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E., (submitted). Preliminary testing of the English Version of the Screener for Intelligence and Learning Disabilities (SCIL) amongst adolescents in Nigeria.

Preliminary testing of the English Version of the Screener for Intelligence and Learning Disabilities (SCIL) amongst adolescents in Nigeria

Eziafakaku Uchechukwu Nwokolo and Glynis H Murphy
University of Kent

Anne-Marie Mensink
Norfolk Community Health and Care NHS Trust

Xavier Moonen
University of Amsterdam

Peter E Langdon
University of Warwick and Coventry and Warwickshire Partnership NHS Trust and Herefordshire and Worcestershire Health and Care NHS Trust

Author Note

Eziafakaku Uchechukwu Nwokolo, Tizard Centre, University of Kent; Glynis H Murphy, Tizard Centre, University of Kent; Anne-Marie Mensink, Norfolk community Health and Care NHS Trust; Xavier Moonen, University of Amsterdam; Peter E Langdon, Centre for Educational Development, Appraisal and Research (CEDAR), and Centre for Mental Health and Wellbeing Research, University of Warwick and Brooklands Hospital, Coventry and Warwickshire Partnership NHS Trust and Herefordshire and Worcestershire Health and Care NHS Trust.

Correspondence concerning this article should be addressed to Eziafakaku Nwokolo, Tizard Centre, University of Kent, Canterbury, CT2 7NF.
E-mail: eun5@kent.ac.uk

Abstract

Background: There are few screening tools for intellectual disabilities that have been developed and used within Africa. The aim of this study was to examine the psychometric properties of the English version of the adolescent Screener for Intelligence and Learning Disabilities (SCIL) when used with Nigerian adolescents.

Method: Two hundred and nine adolescents and young people (aged 11 – 26 years) completed the SCIL and took part in an assessment of their level of general intellectual functioning and adaptive behaviour. Dimension reduction of the SCIL was considered using principal components analysis. Discriminative and convergent validity were examined, along with the sensitivity, specificity, PPV and NPV of the SCIL in identifying those with and without intellectual disabilities.

Results: The SCIL had good internal consistency, discriminative, and convergent validity. Dimension reduction was not necessary. A cut-off score of 10 revealed sensitivity = .74, specificity = .96, PPV = .92 and NPV = .86 for identifying those with an intellectual disability. AUC was .88.

Conclusions: The SCIL has good psychometric properties when used with Nigerian adolescents. Further factor analytic work is needed.

Keywords: adolescent, screening, diagnosis, intellectual disabilities, Nigeria, Africa

Social factors, such as parental decisions, culture, expected outcomes, and beliefs, may undermine the early assessment and diagnosis of mental health or other neurodevelopmental disorders amongst African adolescents (Dogra, 2015; Garland et al., 2004). Thus, adolescents with intellectual disabilities in African countries such as Nigeria are not screened at an early age and often go undiagnosed due to a lack of understanding of the challenges by primary caregivers and some professionals, complicated by the socio-political climate, and lack of adequate tools for screening (Franz, Chambers, von Isenburg, & de Vries, 2017; Nwokolo, Langdon, & Murphy, 2022). The United Nations Children's Fund (UNICEF) projected that 109 million persons under the age of 18 will reside in Nigeria by 2021 (UNICEF, 2014), and further predictions suggest that there will be about 1.1 billion under the age of 18 within Africa by the year 2100 (UNICEF, 2014). Neglecting their mental health and developmental needs, evidenced by the lack of copious research, poses substantial challenges (Kieling et al., 2011; Maxey & Beckert, 2017; Erskine et al., 2017). Based on scant data, the prevalence of adolescent mental illnesses (excluding intellectual disability) stands at 6.7% in Sub-Saharan Africa (Erskine et al., 2017), and the World Health Organization (WHO, 2014) mapped out strategies to address the all-round health concerns of adolescents, which included policies, collaboration between sectors, and data gathering. Given the high comorbidity of mental health disorders in this age group (Jozefiak, Kayed, Rimehaug, Wormdal, Brubakk, & Wichstrøm, 2016; Munir, 2016; Uzun Cıcek, Sarı, & Mercan Isık, 2020), screening for intellectual disability, often included within mental health subdomains, must be included within these initiatives.

A diagnosis of intellectual disability must meet the diagnostic criteria within three domains: a) intellectual functioning, b) adaptive skills, and c) onset in the developmental period, according to the International Classification of Diseases (ICD-11; World Health Organization [WHO], 2020). The administration of complete intelligence tests is usually lengthy, time-consuming, costly and requires trained professionals. Consequently, screening tools were developed and used to save time and cost. Screening helps with the identification of persons who *may* have an intellectual disability and need further assessment. Early identification of adolescents suspected to have an intellectual disability is crucial for adequate classification, diagnosis, and tailoring of the right support and environment for them to thrive (Franz et al., 2017; Matson, Rieske, & Tureck 2011).

In Africa, in countries such as Nigeria, a lack of adequate and validated screening tools and a low level of awareness among parents and professionals have been identified as barriers to assessment (Franz et al., 2017; Nwokolo et al., 2022).

To identify available time- and cost-efficient screening tools for identifying intellectual disabilities with Nigerian adolescents, Nwokolo et al. (2022) conducted a systematic review of such tools examining their cultural appropriateness and psychometric properties. A total of six tools were identified: (1) Hayes Ability Screening Index (HASI), (2) the Learning Disability Screening Questionnaire (LDSQ), (3) the Child and Adolescent Intellectual Disability Screening Questionnaire (CAIDS-Q), (4) the Slosson Intelligence Test (SIT), (5) the Screener for Intelligence and Learning Disabilities (SCIL), and (6) the Quick Test (QT). After assessing the evidence on the six tools, two of them (the SCIL and the CAIDS-Q) were selected for further review by a focus group of Nigerian experts (Nwokolo, Murphy, Mensink, Moonen, & Langdon, submitted). The group of experts examined the face and content validity of both tools. After the focus group consultations in Nigeria, the adolescent version of the Screener for Intelligence and Learning Disabilities (SCIL 14-17) was selected. The group agreed the SCIL had more practical and functional items for screening intellectual disabilities than the CAIDS-Q. The SCIL is a short, 14-item tool for identifying individuals with a level of general intellectual functioning that falls within and below the "very low" range (Nijman et al., 2018). The adolescent version of the SCIL was written in Dutch and translated into English, and minor changes were made following the Nigerian focus group's recommendations regarding the cultural appropriateness of examples in the SCIL for Nigerian adolescents. Lower scores on the SCIL indicate suspicion of possible intellectual disability, with further assessments and diagnosis recommended. Administration of the SCIL is time efficient and does not require any costly or special training.

To ascertain the usefulness of the English version of the adolescent SCIL in Nigeria, as translated and adapted by Nwokolo et al. (submitted), 209 adolescents and young adults were invited to complete the SCIL and undergo measures of their adaptive behaviour and level of general intellectual functioning. This study aimed to a) examine the component structure of the SCIL and reduce dimensions as required, b) examine the internal consistency, discriminative, and convergent validity of the SCIL, c) derive an appropriate cut-off score based upon sensitivity and specificity and d) derive the positive and negative predictive values.

Given these aims, the recruitment of participants from the relevant centres was purposive to allow the inclusion of those with and without an intellectual disability.

Methods

Design

A between-groups design was used with two groups of participants: adolescents and young people thought to have intellectual disabilities and those thought not to have such disabilities.

Participants

The study took place within three geopolitical zones in Nigeria, namely, Enugu, Abuja, and Lagos. In the 'suspected intellectual disability group', an adolescent or young person was eligible to participate in this study if they a) were between 11- and 26 years old, b) were identified by a medical doctor as possibly having an intellectual disability, and/or c) attended a special school or a special day centre. Other participants of the same age, who were thought unlikely to have an intellectual disability, were also recruited.

Participants were recruited from day centres, special schools, child and adolescent mental health care services, local community organisations, places of worship and public advertisement. Initially, 245 adolescents were invited to take part in this study, and 35 declined to participate or did not respond to further attempts to contact them; finally, 210 adolescents (M_{age} = 15.88 years; Mdn_{age} = 15.29 years; SD = 3.69; Min: 10.90 years; Max: 26.96 years; 41% female and 59% male) took part in this study. The age distribution was categorised as follows: 11 – 13-year-olds (n = 76; 36.2%), 14 – 15-year-olds (n = 42; 20%), 16 – 17-year-olds (n = 51; 24.3%) and 18 years and above (n = 41; 19.5%).

Measures

Screener for Intelligence and Learning Disabilities (SCIL)

The SCIL (Nijman et al., 2018; Geijsen et al., 2018) is a standardised 14-item questionnaire used to screen intellectual disabilities and takes 10 – 15 minutes to complete. The measure was developed and used for adolescents in the Netherlands, and it was translated and back translated into English and subject to minor modifications to ensure it was culturally suitable for Nigerians

(see Nwokolo et al., submitted). A further description of the SCIL items is available elsewhere (Geijsen et al., 2018).

Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V)

The WISC-V (Wechsler, 2014) is an individually administered intelligence test for children aged 6 to 16. Administration of the WISC-V takes between 45 to 65 minutes. The child's Full-Scale Intelligence Quotient (IQ) is generated from 7 of the primary subtests – Verbal Comprehension Index (2 subtests), Visual Spatial Index (1 subtest), Fluid Reasoning Index (2 subtests), Working memory Index (1 subtest) and Processing Speed Index (1 subtest).

Wechsler Adult Intelligence Scale, Third Edition (WAIS-III)

The WAIS-III (Wechsler, 1997) is standardised for use with older adolescents and adults and was used to calculate Full-Scale IQ for participants who were older than 16 years. Administration can take around 60 minutes.

Vineland Adaptive Behavior Scales, Third Edition (VABS-3)

The VABS-3 is a standardised semi-structured interview to index adaptive behaviour. It can be completed with an individual, their carer/parent, or a teacher. The carer/parent domain level form was used within this study as it has been recommended for research purposes (Pepperdine & McCrimmon, 2018; Sparrow, Cicchetti, & Saulnier, 2016).

Procedure

A favourable ethical opinion was obtained from the University of Kent, Tizard Centre Ethics Committee, the National Health Research Ethics Committee (NHREC; NHREC/01/01/2007-16/09/2019) and the Federal Neuro-Psychiatric Hospital, Yaba, Lagos, Nigeria (FNPH/HREC/20/09). All participants were provided with written information about participating in this study, including easier-to-read versions. Parental informed consent was sought for those under 18 years of age; and those over 18 consented for themselves. Participants were encouraged to take breaks as needed during testing.

Each participant was invited to complete the SCIL and either the Wechsler Adult Intelligence Scale, 3rd Edition (WAIS-III) or the Wechsler Intelligence Scale for Children, 5th Edition (WISC-V), depending on age. The Vineland Adaptive Behavior Scales, 3rd Edition (VABS-3) were completed with either the parent or caregiver with sufficient developmental knowledge about the

participant. The WAIS-III and the WISC-V were administered by a psychologist, while the VABS-3 and SCIL were administered by a qualified autism service practitioner supervisor trained on the assessments.

Data Analysis

Overview

The Statistical Package for the Social Sciences – IBM SPSS version 26 was used for the analysis. Principal component analysis (PCA) was used to examine the component structure of the English version of the SCIL to determine whether any items should be removed, noting that this was conducted previously using the Dutch version (Geijsen et al., 2018). PCA was performed with a Promax rotation with Kaiser normalisation because a correlation between the components was expected. First, individual items were retained if they correlated at least .30 with another item. Second, the remaining items were then retained if item communalities were at least .30 (Floyd & Widaman, 1995) and the criterion of at least .60 for the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was met. Components with eigenvalue >1 were retained.

Following this, discriminative and convergent validity and internal consistency were examined. Convergent validity was assessed by examining the extent to which the scores on the SCIL correlated with Full-Scale IQ and the VABS-3. Receiver Operating Characteristic (ROC) analyses were used to calculate the Area under the Curve (AUC) to examine how well the SCIL identified participants with and without an intellectual disability, with reference to sensitivity and specificity to identify optimal cut-offs to examine discriminative validity. Internal consistency was considered by calculating Cronbach's α. The positive predictive value (PPV) and the negative predictive value (NPV) were determined using the cross-validation of the classification results.

IQ and VABS cut-offs for Intellectual Disabilities

Participants were classified as having an intellectual disability using ICD-11 (WHO, 2020) diagnostic criteria for disorders of intellectual development. Specifically, those with a level of general intellectual functioning and adaptive behaviour that fell at least two standard deviations or more below the mean were classified as having an intellectual disability. However, there is evidence (Wicherts, Dolan, & van der Maas, 2010; Ani & Grantham-McGregor, 1998; Nenty &

Dinero, 1981; Ashem & Janes, 1978, Nwanze & Okeowo, 1980 and Fahrmeier, 1975) to show that the average IQ of Africans on Western IQ tests is at least one standard deviation lower than Western normative data. Questions and discussions around the validity and cultural appropriateness of IQ tests such as the WISC and WAIS developed in the West (predominantly with White populations) and used in Africa or with the Black populations continue to occur (Kamin, 2006; Rushton & Jensen; 2006; Bakare, Ubochi, Okoroikpa, Aguocha & Ebigbo, 2009; Shuttleworth-Edwards, Donnelly, Reid & Radloff, 2004; Shuttleworth-Edwards, Kemp, Rust, Muirhead, Hartman & Radloff, 2004). Recognising that environment, culture, exposure to Western cultures, level and quality of education, and various other factors contribute to an estimate of the level of general intellectual functioning (Bakare et al., 2009; Shuttleworth-Edwards, Kemp, Rust et al., 2004), it would be "illogical to use scores on such tests to infer the level of innate ability possessed by people in non-Western cultures or as a basis for making judgmental statements of a superior-subordinate nature about their performance" (Bakare, 1972, p.362). However, Wicherts et al. (2010) systematically reviewed published empirical data on various Western IQ tests with Western norms regarding the performance of African populations. The tests included the Kaufman-Assessment Battery for Children (K-ABC), the Wechsler scales (WAIS & WISC), the Draw-A-Man (DAM) test, and several others, with the goal of estimating the average IQ of samples of normal and healthy Africans. Their results showed that the average IQ of Africans was approximately 82 when compared to UK norms.

Regarding the Vineland Adaptive Behavior Scales (VABS), several studies have compared the VABS to measures of adaptive behaviour that were developed in African countries (du Toit, Van der Linde & Swanepoel, 2021a; du Toit, Van der Linde & Swanepoel, 2021b; de Beer, Krüger, Van der Linde, Eccles & Graham, 2020; Allen et al., 2014). Unfortunately, the authors of these studies did not report the data captured using the VABS to allow for conclusions about how well this instrument performs when used within Africa (Beer et al. 2020; du Toit et a. 2021a, 2021b; Allen et al. 2014). Douglas (2017) concluded that the VABS can identify persons with intellectual disability under 22-years of age amongst sexual abuse victims in South Africa but did not report their actual VABS data making it difficult to examine how well the VABS performed.

Considering the issues mentioned above and our work's cultural sensitivity, we made use of two different IQ cut-offs for identifying participants with an intellectual disability and completed our

analysis twice. Initially, we used the established, **unadjusted**, Western criterion of FSIQ <70 (i.e., 2 SD below 100); then, secondly, the **FSIQ cut-off was adjusted** based on the work of Wicherts et al. (2010) to FSIQ score <52 (i.e., 2 SD below 82). There is a lack evidence about the performance of the VABS-3 when used to identify intellectual disability within Africa. Considering this, we opted to retain and use the established cut-off of <70 (i.e., 2 SD below 100) for the VABS for deciding whether a participant had an intellectual disability for our analysis using both the adjusted and unadjusted Full Scale IQ cut-off. In a small number of cases (n = 8), the VABS-3 Composite score was above 70 while the FSIQ was below 70: in these cases, they were allocated to the 'no intellectual disability, unadjusted cut-off' group. Also, in a further small number of cases (n = 9), the FSIQ was above 52 while the VABS-3 Composite score was below 70: in these cases, they were allocated to 'no intellectual disability, adjusted cut-off' group.

Missing Data

Thirty-seven participants (30 with WISC-V, 7 with WAIS-III) could not complete the assessment of their level of general intellectual functioning, as they were scoring **below the floor of the test**, possibly due to the degree of intellectual disability, while one of these participants was also unable to score on the SCIL and the VABS-3. Therefore, thirty-seven participants were excluded from analyses using the Weschler scales, while 1 participant was excluded from the analysis using VABS Composite scores and SCIL analyses.

Results

Descriptive Statistics

SCIL scores for the whole group of participants ranged from 0 to 28 points, M = 14.01; Mdn = 16.00; SD = 9.4. SCIL scores did not differ between the sexes, $t_{(207)} = -1.818$, $p = .07$. Descriptive statistics for the SCIL by age group for the whole sample are in Table 1. There was no significant correlation between age and the SCIL scores $r_{(209)} = .13$, $p = .06$. Likewise, Tables 2 and 3 show the mean and SD for VABS-3 Composite scores for the whole group and for FSIQ scores for the whole group. There was significant correlation between the VABS-3 Composite scores and the FSIQ scores $\mathbf{r}_{(173)} = .74$, $p < .001$ (excluding the 37 who scored below the floor on the Weschler Scales).

When the initial criterion (FSIQ <70) for very low intellectual functioning was applied to the whole sample, 68.57% (N = 144) were classified as having an intellectual disability, 30.95% (N = 65) as not having an intellectual disability and 0.5% (N = 1) missing. However, applying the criterion of FSIQ <52 to the whole sample, 70% (N = 147) of participants were reclassified as not having an intellectual disability and 29.52% (N = 62) with an intellectual disability. The VABS-3 Composite scores criterion of <70 classified 69.05% (N = 145) as not having an intellectual disability and 30.48% (N = 64) with an intellectual disability, and 0.5% (N = 1) missing.

It is important to note that these figures do NOT give the prevalence of intellectual disability in Nigeria since the two groups were specifically chosen so as to represent likely intellectual disability and unlikely intellectual disability. As previously mentioned, 37 participants did not complete the FSIQ assessment as they were scoring below the floor of the test, possibly due to the level of their intellectual functioning and were subsequently excluded from these analyses of FSIQ.

Table 1 About Here

Table 2 About Here

Table 3 About Here

Thirty-seven percent of participants were initially categorised as having an intellectual disability as they had a Full-Scale IQ and VABS Composite Score of <70 (Table 4). Adjusting the Full-Scale IQ cut-off to <52 resulted in 16% of our participants being categorised as having an intellectual disability (Table 5).

Table 4 About Here

Finally, the participants were divided again, this time using the adjusted cut-off of IQ 52, so as to give the mean scores and SD for the 'Intellectual Disabilities, adjusted cut-off' and 'no Intellectual Disabilities, adjusted cut-off' groups, as described above in the Method section. Table 5 gives the resulting FSIQ scores for these two groups.

Table 5 About Here

Principal Component Analysis (PCA)

The correlation matrix revealed that all items of the SCIL correlated at $r > .30$ with more than one other item; thus, all 14 items of the SCIL were initially retained. Multicollinearity criteria ($r > .8$) analysis showed that all correlations were less than .8. The KMO was .94, and all KMOs for individual items were greater than .87, which is well above the acceptable limit of .5 (Field, 2013; Kaiser, 1974). Bartlett's test of sphericity was significant, $\chi 2(91) = 1947.49$, $p < .001$, which implies that the items are correlated. Therefore, all items were retained. The scree plot was slightly ambiguous, showing inflexions that would justify retaining three factors. However, two factors had eigenvalues >1 and explained 63.56% of the variance. The Pattern Matrix and Structure Matrix are in Table 6. A full description of the SCIL items is in Appendix 1. The 9 items (1, 2, 3, 4, 8, 9, 10, 12, 13) on component 1 were judged to relate to Education, Social

Contacts, and Comprehension, while the 5 items (5, 6, 7, 11, 14) on component 2 were judged to relate to Arithmetic and Numbers.

Table 6 About Here

Internal Consistency

The overall Cronbach's alpha value of .94 indicated the high internal consistency of the SCIL items. Component 1 had a value of .92, while component 2 was .84, which were both high.

Convergent Validity

There was a significant positive correlation between the total SCIL score and level of general intellectual functioning (judged by the FSIQ Score), $r_{(173)} = .81$, $p < 0.001$ (excluding the 37 who scored below the floor of the test), indicating a large effect size (Cohen, 1992). The correlation between the total SCIL score and Full-Scale IQ estimated using the WAIS-III, $r_{(52)} = .86$, $p < 0.001$, and the WISC-V, $r_{(121)} = .79$, $p < 0.001$, was high. Similarly, there was a significant positive correlation between the total SCIL score and VABS-3 Composite Score, $r_{(209)} = .84$, $p < 0.001$ for all the participants.

Sensitivity and Specificity *(FSIQ and VABS-3 Composite Scores <70)*

At the suggested SCIL cut-off score of 15 (Nijman et al., 2018), the AUC was .81, $p < 0.001$, 95% CI [.75, .86], sensitivity = 1 and specificity = .55, applicable to the entire sample. However, exploring lower cut-off scores of 10, 11 and 12 did not improve the specificity values (.44, .52, and .54, respectively), although sensitivity remained at 1, implying that these cut-off scores were too low. For this study, there was no difference in categorising participants with likely or unlikely intellectual disability using the cut-off scores of 13 and 14 versus the suggested 15. As such, the sensitivity and specificity using 13 or 14 were not explored. Using 15 as the cut-off, the AUC across the age groups was moderate: 11 – 13-year-olds, AUC = .85, $p < 0.001$, 95% CI [.76, .93], $N = 76$, 14 – 15-year-olds, AUC = .74, $p < 0.001$, 95% CI [.60, .89], $N = 42$, 16 – 17-

year-olds, AUC = .79, $p < 0.001$, 95% CI [.67, .91], $N = 51$, 18 years and above, AUC =.83, $p < 0.001$, 95% CI [.71, .96], $N = 40$.

Positive Predictive Value (PPV) and Negative Predictive Value (NPV) ***(FSIQ and VABS-3 Composite Scores <70)***

At the suggested SCIL cut-off score of 15, (Nijman et al., 2018), the PPV was .62 and the NPV was 1.

Sensitivity and Specificity ***(FSIQ <52 and VABS-3 Composite Scores <70)***

When the adjusted FSIQ (i.e., with the cut-off of 52) and VABS-3 Composite scores (i.e., 2 SDs below 100) were used, using the suggested SCIL cut-off score of 15 (Nijman et al., 2018), sensitivity = 1 and specificity = .78 for the entire sample. Both measurement properties met the minimum standard (Glascoe, 2005). The AUC was .89, $p < 0.001$, 95% CI [.85, .93]. Also, using 15 as the cut-off, the AUC across the age groups was large: 11 – 13-year-olds, AUC = .83, $p < 0.001$, 95% CI [.75, .92], $N = 76$, 14 – 15-year-olds, AUC = .95, $p < 0.001$, 95% CI [.88, 1.02], $N = 42$, 16 – 17-year-olds, AUC = .90, $p < 0.001$, 95% CI [.82, .98], $N = 51$, 18 years and above, AUC = .91, $p < 0.001$, 95% CI [.83, 1.00], $N = 40$. To determine the best cut-off score, rather than selecting an arbitrary figure, lower cut-offs were explored by stepwise reduction. Lowering the cut-off score to 10, 11 and 12 improved the values; however, a cut-off score of 10 gave the best result. A cut-off score of 13 did not yield different results from using 15 for this study and was not explored further. The AUC indicates the discriminative ability of the SCIL; a perfect tool would have an AUC of 1, and the AUCs in this study ranged from .83 to .95, indicating good discriminative validity. The AUC, PPV, NPV, sensitivity and specificity associated with each cut-off score are shown in Table 7 – figures in bold indicate measurement properties at the applicable cut-off scores.

Positive Predictive Value (PPV) and Negative Predictive Value (NPV) ***(FSIQ <52 and VABS-3 Composite Scores <70)***

Using the cut-off of 10, the SCIL had a PPV = .66 and NPV = 1, and with a cut-off of 11, PPV = .65 and NPV = 1, while at a cut-off of 12, PPV = .61 and NPV = 1. These show that the SCIL can correctly identify those with and without intellectual disabilities. However, the most suitable cut-off was determined to be 10.

Table7 About Here

Discussion and Conclusion

The aims of this study were to examine a) the component structure of the SCIL, b) the internal consistency, discriminative, and convergent validity of the SCIL, c) the likely appropriate cut-off score based upon sensitivity and specificity, and d) the positive and negative predictive values. Early identification of intellectual disabilities is essential for significant progress in intervention, educational support, and policy (Luckasson & Schalock, 2013; Schalock & Luckasson, 2013), but many adolescents in Nigeria have gone undiagnosed in early childhood (Franz et al., 2017).

Component Structure

Following an examination of the component structure of the SCIL, all 14 items were retained. Two components were derived, which we labelled (1) Education, Social Contacts, and Comprehension, and (2) Arithmetic and Numbers. Geijsen et al. (2018) previously used PCA to examine the component structure of the SCIL and retained 4 components. This disparity in findings may be due to the sampling environment, sample characteristics and the use of a different version of the SCIL. Geijsen et al. (2018) study was conducted with Dutch participants aged between 18 and 63 years in police custody, whereas participants for this study were Nigerian adolescents aged between 11 and 26 years from schools, CAMHS and the public. Also, this study employed the adolescent SCIL, whereas Geijsen et al. (2018) used the adult version of the SCIL.

Internal Consistency and Convergent Validity

The result indicated that the internal consistency of the SCIL was excellent. Convergent validity measures the relationship between two related constructs; in this study, it was the relationship between the SCIL scores, FSIQ and VABS-3 Composite scores. For this, positive relationships were found between FSIQ, the VABS-3 Composite scores and SCIL scores.

Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value (FSIQ and VABS-3 Composite scores <70)

ROC analyses were conducted separately for the entire sample and for each age group. The specificity did not meet the minimum criteria of 70% at the suggested cut-off of 15 (Nijman et al., 2018), while sensitivity met the criteria (Glascoe, 2005). At the suggested cut-off of 15, about 55% of participants would have been excluded as not having a possible intellectual disability. Lowering the cut-off to 10, 11 and 12 did not improve the results. The resultant low specificity at FSIQ <70 and VABS-3 Composite score <70 for the SCIL is possibly due to the comparison of the mean IQ of our study participants to the Wechsler scales based on UK norms. These cut-off scores are not likely to represent deficits in cognitive abilities, given that the FSIQ and VABS-3 Composite scores used were based on the UK norm. Psychometric assessments in cross-cultural environments are problematic, considering the role that culture, beliefs, language, ethnicity, exposure, and quality of education play in test-taking.

Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value (FSIQ <52 and VABS-3 Composite Scores <70)

When using adjusted norms (FSIQ <52 and VABS-3 Composite <70), ROC analyses were conducted separately for the entire sample and for each age group, and in all cases, the sensitivity, specificity, and NPV met the minimum acceptable criteria of 70% accuracy for screening tools (Glascoe, 2005) while the PPV was within an acceptable range. An entire sample cut-off score of 10 was deemed appropriate for this study. Specific cut-offs were derived for the different age groups: 11 – 13 years (10), 14 – 15 years (10), 16 – 17 years (11), and 18 years and above (11). The sensitivity and specificity for all met the minimum criteria (70). In assigning these cut-offs, consideration was given to the distinctions between sensitivity and PPV and between specificity and NPV in a screening and clinical context (Trevethan, 2017; Akobeng, 2007). Classifying participants based solely on sensitivity and specificity values differ from classifying them in combination with the PPV and NPV. PPV and NPV are influenced by the condition's prevalence and depend on the population being investigated. Thus, in selecting 10 as the cut-off score, a combination of the PPV, NPV, sensitivity and specificity was used. Nijman et al. (2018) previously expressed concerns about the appropriateness of using the SCIL with the younger population (12 to 13-year-olds). However, our findings indicated that the SCIL could identify younger participants with and without intellectual disabilities. It is important to note that

people attain developmental milestones at different times; having different cut-offs for the different age ranges is useful, noting that the cut-off score associated with each age group was near the overall cut-off of 10.

Limitations

There are limitations to this study. First, the SCIL was used with a sample of adolescents drawn mainly from large urban cities in Nigeria, which may have introduced some bias, as this sample is not representative of the whole population of Nigeria. Secondly, the sample size was reduced by 37 due to some individual's FSIQ scores being below the floor of the test. However, our sample size remained appropriately large for our chosen analyses. Thirdly, the test-retest reliability of the SCIL was not examined. Fourthly, the VABS-3, has not been validated for use in Africa or Nigeria, thus, there is uncertainty about the cut-off score that should be used to identify those with intellectual disabilities.

The degree of uncertainty regarding the cultural sensitivity of the VABS-3 is currently unknown; however, in the absence of an alternative measure for adaptive behaviour and considering the observations from Tan, Reich, Hart, Thuma, & Grigorenko (2014) and other VABS studies previously mentioned, we utilised the VABS-3. Further VABS-3 validation work in Nigeria and Africa is needed. Lastly, the adjustments to the mean FSIQ used to identify participants with intellectual disabilities may be potentially problematic and sensitive in nature, which we wish to recognise. However, as previously mentioned, there are studies to support the adjustments. Despite these limitations, our findings indicated that the SCIL is a valid screening tool for intellectual disability in Nigerian adolescents.

As previously mentioned, there are limitations to the use of sensitivity and specificity, which can be overcome by also examining the PPV and NPV. Trevethan (2017, p.4) considered that "… sensitivity and specificity indicate the concordance of a test with respect to a chosen referent, while PPV and NPV, respectively, indicate the likelihood that a test can successfully identify whether people do or do not have a target condition, based on their test results." Our results demonstrated that the English translation of the SCIL can be used for screening for intellectual disabilities in the Nigerian adolescent population. The reservations regarding its usefulness in younger adolescents (below 14 years) raised by Nijman et al. (2018) can be overcome by examining the PPV and NPV in addition to the sensitivity and specificity for the specific age

groups. For this study, sensitivity, specificity, PPV and NPV were adequate (1, 0.70, 0.53 & 1 respectively) for those below 14 years.

Conclusions

Our findings indicated that the English version of the adolescent SCIL has good construct, convergent and discriminative validity. The SCIL can offer a valuable means of identifying adolescents likely to have an intellectual disability to facilitate intervention at an earlier stage (Franz et al., 2017), provide targeted support (Kieling et al., 2011), and help ensure referrals for further diagnosis. As a simple and quick screening tool, further research utilising the English version of the adolescent SCIL in more Nigerian cities and in other English-speaking African populations is recommended.

Financial Disclosure

None

Funding Source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors do not have any conflict of interest to declare.
Xavier Moonen is one of the developers of the Dutch version of the SCIL. He receives an annual financial contribution from the publisher for, among other things, further research.

References

Akobeng, A. K. (2007). Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatrica*, 96(3), 338-341.

Alice, M. (2018). Imputing missing data with R; MICE package. Retrieved from https://datascienceplus.com/imputing-missing-data-with-r-mice-package/ Accessed 14 February 2022.

Allen, A. B., Finestone, M., Eloff, I., Sipsma, H., Makin, J., Triplett, K., ... & Forsyth, B. W. (2014). The role of parenting in affecting the behavior and adaptive functioning of young children of HIV-infected mothers in South Africa. *AIDS and Behavior*, *18*, 605-616.

Ani, C. C., & Grantham-McGregor, S. (1998). Family and personal characteristics of aggressive Nigerian boys: Differences from and similarities with Western findings. *Journal of Adolescent Health*, *23*(5), 311-317.

Ashem, B., & Janes, M. D. (1978). Deleterious effects of chronic undernutrition on cognitive abilities. *Journal of Child Psychology and Psychiatry*, *19*, 23-31.

Bakare, C. G. M. (1972). Social-class differences in the performance of Nigerian children on the Draw-a-Man test. In L. J. Cronbach, & P. J. Drenth (Eds.), *Mental tests and cultural adaptation* (pp. 355–363). The Netherlands: Mouton: The Hague.

Bakare, M. O., Ubochi, V. N., Okoroikpa, I. N., Aguocha, C. M., & Ebigbo, P. O. (2009). Agreement between clinicians' and care givers' assessment of intelligence in Nigerian children with intellectual disability: 'ratio IQ' as a viable option in the absence of standardized 'deviance IQ' tests in sub-Saharan Africa. *Behavioral and Brain Functions*, *5*(1), 1-7.

Bland, M. (2015). An introduction to medical statistics. Fourth Edition. *Types of missing data*. Oxford University Press. Retrieved from https://www-users.york.ac.uk/~mb55/intro/typemiss4.htm Accessed 16 February 2022.

Boat, T. F., Wu, J. T., & National Academies of Sciences, Engineering, and Medicine. (2015). Chapter 9. Clinical characteristics of intellectual disabilities. *Mental disorders and disabilities among low-income children*. Washington (DC): National Academies Press (US). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK332877/ Accessed 7 February 2022.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

de Beer, C. C., Krüger, E., Van der Linde, J., Eccles, R., & Graham, M. A. (2020). Developmental outcomes of HIV-exposed infants in a low-income South African context. *African Health Sciences*, *20*(4), 1734-41.

Dogra, N. (2015). Child and adolescent psychiatry, principles of. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)* (pp. 383-390). Oxford: Elsevier.

Douglas, G. K. (2017). *Use of the Vineland Adaptive Behavior Scale in the assessment of intellectually disabled complainants in sexual abuse cases in the Western Cape* (Doctoral dissertation, Stellenbosch: Stellenbosch University).

du Toit, M., Van der Linde, J., & Swanepoel, D. W. (2021a). Early childhood development risks and protective factors in vulnerable preschool children from low-income communities in South Africa. *Journal of Community Health*, *46*, 304-312.

du Toit, M. N., Van der Linde, J., & Swanepoel, D. W. (2021b). mHealth developmental screening for preschool children in low-income communities. *Journal of Child Health Care*, *25*(4), 573-586.

Erskine, H., Baxter, A., Patton, G., Moffitt, T., Patel, V., Whiteford, H., & Scott, J. (2017). The global coverage of prevalence data for mental disorders in children and adolescents. *Epidemiology and Psychiatric Sciences*, *26*(4), 395-402.

Fahrmeier, E. D. (1975). The Effect of School Attendance on Intellectual Development in Northern Nigeria. Child Development, 46(1), 281–285.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (Fourth ed.) London. Sage.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286.

Franz, L., Chambers, N., von Isenburg, M., & de Vries, P. J. (2017). Autism spectrum disorder in sub-Saharan Africa: A comprehensive scoping review. *Autism Research*, 10(5), 723-749.

Garland, A. F., Lewczyk-Boxmeyer, C. M., Gabayan, E. N., & Hawley, K. M. (2004). Multiple stakeholder agreement on desired outcomes for adolescents' mental health services. *Psychiatric Services, 55*(6), 671-676.

Geijsen, K., Kop, N., & de Ruiter, C. (2018). Screening for intellectual disability in Dutch police suspects. *Journal of Investigative Psychology and Offender Profiling, 15*(2), 200-214.

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(3), 173-179.

Goldberg, M. R., Dill, C. A., Shin, J. Y., & Nhan, N. V. (2009). Reliability and validity of the Vietnamese Vineland Adaptive Behavior Scales with preschool-age children. *Research in Developmental Disabilities*, *30*(3), 592-602.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*(2), 265.

Hessl, D., Nguyen, D. V., Green, C., Chavez, A., Tassone, F., Hagerman, R. J., Senturk, D., Schneider, A., Lightbody, A., Reiss, A. L. & Hall, S. (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: The case of fragile X syndrome. *Journal of Neurodevelopmental Disorders, 1*(1), 33-45.

Jozefiak, T., Kayed, N. S., Rimehaug, T., Wormdal, A. K., Brubakk, A. M., & Wichstrøm, L. (2016). Prevalence and comorbidity of mental disorders among adolescents living in residential youth care. *European Child & Adolescent Psychiatry*, *25*(1), 33-47.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36.

Kamin, L. J. (2006). African IQ and mental retardation. *South African Journal of Psychology*, *36*(1), 1-9.

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *64*(5), 402-406.

Katsiana, A., Stalikas, A., Kokkaris, P., Galanakis, M., Georgiou, K., & Strimpakos, N. (2022). Cross-Cultural Adaptation and Psychometric Properties of the Greek Vineland Adaptive Behavior Scales: Parent/Caregiver Rating Form (VABS II-Gr). *Psychology*, *13*(13), 1850-1864.

Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., Rohde, L. A., Srinath, S., Ulkuer, N. & Rahman, A. (2011). Child and adolescent mental health worldwide: evidence for action. *The Lancet*, 378(9801), 1515-1525.

Kumar, R., Shankar, K., Kush, V., Kumar, C., Bhave, A., & Agarwal, V. (2016). Adaptation: Vineland Adaptive Behavior Scale for 3–9 year-old Indian children. *International Journal on Disability and Human Development*, *15*(1), 49-55.

Lalkhen, A. G., & McCluskey, A. (2008). Clinical tests: Sensitivity and specificity. *Continuing Education in Anaesthesia Critical Care & Pain*, *8*(6), 221-223.

Lasko, T. A., Bhagwat, J. G., Zou, K. H., & Ohno-Machado, L. (2005). The use of receiver operating characteristic curves in biomedical informatics. *Journal of Biomedical Informatics*, *38*(5), 404-415.

Luckasson, R., & Schalock, R. L. (2013). What's at stake in the lives of people with intellectual disability? Part II: Recommendations for naming, defining, diagnosing, classifying, and planning supports. *Intellectual and Developmental Disabilities*, 51(2), 94-101.

Lynn, R. (2015). *Race differences in intelligence: An evolutionary analysis.* (Second ed.) Washington Summit Publishers. Retrieved from https://www.intelligence-humaine.com/wp-content/uploads/2019/03/Race-Differences-in-Intelligence-second-edition-2015-1.pdf Accessed 21 November 2022.

Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, *110*, 63-73.

Matson, J. L., Rieske, R. D., & Tureck, K. (2011). Additional considerations for the early detection and diagnosis of autism: Review of available instruments. *Research in Autism Spectrum Disorders*, *5*(4), 1319-1326.

Maxey, M., & Beckert, T. E. (2017). Adolescents with disabilities. *Adolescent Research Review*, 2(2), 59-75.

McKenzie, K., Paxton, D., Murray, G., Milanesi, P., & Murray, A. L. (2012). The evaluation of a screening tool for children with an intellectual disability: The Child and Adolescent Intellectual Disability Screening Questionnaire. *Research in Developmental Disabilities, 33*(4), 1068-1075.

Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2018). COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). *User manual.* *https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf* Accessed 28 November 2018.

Munir, K. M. (2016). The co-occurrence of mental disorders in children and adolescents with intellectual disability/intellectual developmental disorder. *Current Opinion in Psychiatry*, *29*(2), 95.

National Research Council (US) Committee on Disability Determination for Mental Retardation. (2002). Chapter 3, The role of intellectual assessment. In Reschly, D. J., Myers, T. G., Hartel, C. R., (Ed.), *Determining eligibility for social security benefits*. Washington (DC): National Academies Press (US). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK207539/ Accessed 7 February 2022.

Nenty, H. J., & Dinero, T. E. (1981). A cross-cultural analysis of the fairness of the Cattell Culture Fair Intelligence Test using the Rasch model. *Applied Psychological Measurement, 5*(3), 355-368.

Nijman, H., Kaal, H., van Scheppingen, L., & Moonen, X. (2018). Development and testing of a Screener for Intelligence and Learning Disabilities (SCIL). *Journal of Applied Research in Intellectual Disabilities, 31*(1), e59-e67.

Nwokolo, E. U., Langdon, P. E., & Murphy, G. H. (2022). Screening for intellectual disabilities and/or autism amongst older children and young adults: A systematic review of tools for use in Africa. *Review Journal of Autism and Developmental Disorders*. Advance online publication.

Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E., (submitted). Using the consensus group method to select the best screening tools for autism and intellectual disability for use with Nigerian adolescents. *Journal of Policy & Practice in Intellectual Disabilities.*

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology, 56*(1), 45-50.

Pepperdine, C. R., & McCrimmon, A. W. (2018). Test review: Vineland Adaptive Behavior Scales, Third Edition (Vineland-3) by Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. *Canadian Journal of School Psychology*, *33*(2), 157-163.

Rushton, J. P., & Jensen, A. R. (2006). The totality of available evidence shows the race IQ gap still remains. *Psychological Science-Cambridge-*, *17*(10), 921.

Schalock, R. L. & Luckasson, R. A. (2013) What's at stake in the lives of people with intellectual disability? Part I: The power of naming, defining, diagnosing, classifying, and planning supports. *Intellectual and Developmental Disabilities*, 51(2), 86-93.

Shuttleworth-Edwards, B. A., Donnelly, J. R., Martin, I. R., & Radloff, E. S. (2004). A cross-cultural study with culture fair normative indications on WAIS-III Digit Symbol—Incidental learning. *Journal of Clinical and Experimental Neuropsychology, 26*(7), 921-932.

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G., Hartman, N. P., & Radloff, S. E. (2004). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAIS-III test performance. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 903-920.

Sinharay, S., Stern, H. S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods, 6*(4), 317.

Sparrow, S., Cicchetti, D., & Saulnier, C. (2016). *Vineland Adaptive Behavior Scales–Third Edition (Vineland-3)*. Circle Pines, MN: American Guidance Service.

Streiner, D. L., & Cairney, J. (2007). What's under the ROC? an introduction to receiver operating characteristics curves. *The Canadian Journal of Psychiatry, 52*(2), 121-128.

Tan, M., Reich, J., Hart, L., Thuma, P. E., & Grigorenko, E. L. (2014). Examining the specific effects of context on adaptive behavior and achievement in a rural African community: Six case studies from rural areas of Southern Province, Zambia. *Journal of autism and developmental disorders*, *44*, 271-282.

Tassé, M. J., Luckasson, R., & Schalock, R. L. (2016). The relation between intellectual functioning and adaptive behavior in the diagnosis of intellectual disability. *Intellectual and Developmental Disabilities*, *54*(6), 381-390.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34-42.

Touil, N., Riche, B., Des Portes, V., Mardirossian, S., Gaillard, S., Rabilloud, M., ... & Sonie, S. (2021). A French adaptation of the Vineland Adaptive Behavior Scales VABS-II. *European Psychiatry*, *64*(S1), S155-S156.

Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health*, *5*, 307.

Tucker, L. R., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika, 34*(4), 421-459.

UNICEF. (2014). Generation 2030 Africa: Child demographics in Africa. Division of Data. *Research Policy.* Retrieved from https://data.unicef.org/resources/generation-2030-africa-child-demographics-in-africa/ Accessed 4 February 2022.

Uzun Cıcek, A., Sarı, S. A., & Mercan Isık, C. (2020). Sociodemographic characteristics, risk factors, and prevalence of comorbidity among children and adolescents with intellectual disability: A cross-sectional study. *Journal of Mental Health Research in Intellectual Disabilities, 13*(2), 66-85.

Vogt, D. S., King, D. W., & King, L. A. (2004). Focus groups in psychological assessment: Enhancing content validity by consulting members of the target population. *Psychological Assessment, 16*(3), 231.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children (5th ed.)*. Bloomington, MN: NCS Pearson.

Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. (2010). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence, 38*(1), 1-20.

Wong, H. B., & Lim, G. H. (2011). Measures of diagnostic accuracy: Sensitivity, specificity, PPV and NPV. *Proceedings of Singapore Healthcare, 20*(4), 316-318.

World Health Organization. (2014). *Health for the World's Adolescents: A Second Chance in the Second Decade: Summary*. Retrieved from https://apps.who.int/iris/handle/10665/112750. Accessed 4 February 2022.

World Health Organization. (2020). *International Statistical Classification of Diseases and*

Related Health Problems (ICD) (11th Edition). https://icd.who.int/browse11/l-m/en. Accessed 12 November 2020.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432.

Appendix 23 – Validation of the SCQ – Submitted Manuscript

A version of this paper was submitted to the journal Autism Research, in March 2023, as Nwokolo, E. U., Murphy, G. H., & Langdon, P. E. Validation of the Social Communication Questionnaire (SCQ) amongst Nigerian adolescents.

Validation of the Social Communication Questionnaire (SCQ) amongst Nigerian adolescents

Eziafakaku Uchechukwu Nwokolo and Glynis H Murphy
University of Kent

Peter E Langdon
University of Warwick and Coventry and Warwickshire Partnership NHS Trust

Author Note

Eziafakaku Uchechukwu Nwokolo, Tizard Centre, University of Kent; Glynis H Murphy, Tizard Centre, University of Kent; Peter E Langdon, Centre for Educational Development, Appraisal and Research (CEDAR), and Centre for Mental Health and Wellbeing Research, University of Warwick and Brooklands Hospital, Coventry and Warwickshire Partnership NHS Trust.

Correspondence concerning this article should be addressed to Eziafakaku Nwokolo, Tizard Centre, University of Kent, Canterbury, CT2 7NF.
E-mail: eun5@kent.ac.uk

Abstract

Few autism spectrum disorder (ASD) screening tools have been developed and validated in Africa. This study aimed to examine the psychometric properties of the Social Communication Questionnaire (SCQ) when used with Nigerian adolescents. Parents and caregivers of two hundred and five adolescents completed the SCQ Lifetime form while the adolescents were assessed for ASD using the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2). Factor structure and convergent and discriminative validity were examined, along with the sensitivity and specificity of the SCQ in identifying participants with an autism spectrum disorder. The SCQ had good internal consistency, discriminative, and convergent validity. A cut-off score of 10 revealed sensitivity = .75 and specificity = .91 for the identification of autism spectrum disorder. AUC was .91. The results of this study provide evidence to support the retention of the original four factors of the SCQ. The SCQ has good psychometric properties when used with Nigerian adolescents.

Lay Summary

We aimed to validate the English version of the Social Communication Questionnaire (SCQ) amongst Nigerian adolescents aged 11–26 years as a screening tool for identifying individual's who may be at risk of autism spectrum disorder (ASD). The SCQ showed good psychometric properties. We also provided different cut-offs to identify ASD cases for different age groups.

Keywords: adolescent, screening, diagnosis, autism, Nigeria, Africa

The diagnosis of autism spectrum disorder (ASD), a condition characterised by restricted and repetitive behaviours and social and communication deficits, has become increasingly common (Wing & Potter, 2002). With no known cure, a series of studies have found that lifetime costs for individuals with ASD can run into hundreds of thousands of dollars (Horlin, Falkmer, Parsons, Albrecht & Falkmer, 2014; Penner, Rayar, Bashir, Roberts, Hancock-Howard & Coyte, 2015; Sampaio, Feldman, Lavelle & Skokauskas, 2021; Rosenberg, Landa, Law, Stuart & Law 2011). However, early diagnosis and intervention have been shown to produce progress in independent functioning levels, development rate and access to effective services (James & Smith, 2020; Delehanty, Lee, Hooker, Cortese & Woods, 2020; Nadel & Poss, 2007). Nevertheless, appropriate and prompt diagnoses are crucial for accessing such intervention services early in life to capitalise on these gains.

Screening and diagnosis of ASD are feasible in very young children and are recommended as best practice; however, this has not been the norm in Nigeria, as most individuals with ASD are not diagnosed until after five years of age, and many are never diagnosed (Franz, Chambers, von Isenburg & de Vries, 2017; Bello-Mojeed, Bakare & Munir, 2013; Bello-Mojeed, Omigbodun, Bakare & Adewuya, 2017). Different factors, such as low level of awareness, limited availability of qualified professionals, cultural differences, and access to standardised tools, affect the early assessment and diagnosis of developmental disorders such as autism spectrum disorder (ASD) amongst African adolescents (Franz et al., 2017; Bello-Mojeed et al., 2013; Bello-Mojeed et al., 2017; Burkett, Morris, Manning-Courtney, Anthony & Shambley-Ebron, 2015). Therefore, adolescents with ASD in African countries such as Nigeria often go undiagnosed.

Currently, ASD diagnoses in Nigeria do not include the use of a standardised schedule (Oshodi et al., 2017; Bakare et al., 2022). Whereas a clinical assessment of autism spectrum disorder can be given based on history taking, observation and use of the DSM-5 criteria by healthcare professionals, a confirmatory diagnosis using an acceptable gold standard schedule is required for better certainty (Zeidan, Fombonne, Scorah, Ibrahim, Durkin, Saxena, et al., 2022; McCarty & Frye, 2020). The cost of acquiring such a tool is not only prohibitive, but the administration also requires trained professionals, both of which are scarce in Nigeria (Abubakar, Ssewanyana, & Newton, 2016). Therefore, level two screening tools can be used to save time and cost, and such screening would help with the immediate identification of individuals at risk of ASD. In the

African context, in countries such as Nigeria, however, limited availability of age-appropriate screening and validated screening tools, as well as low levels of awareness among parents and professionals, have been identified as barriers to assessment (Franz et al., 2017; Nwokolo, Langdon & Murphy, 2022).

A systematic review was conducted to identify available brief and cost-efficient screening tools for use with Nigerian adolescents (Nwokolo et al., 2022), aiming to judge their cultural appropriateness and assess the psychometric properties of available tools. A total of 12 screening tools for ASD were identified through that review. The tools were the Social Communication Questionnaire (SCQ), the Childhood Autism Rating Scale (CARS), the Child Behavior Checklist (CBCL), the Pervasive Developmental Disorder in Mentally Retarded Persons Scale (PDD-MRS), the Autism Screening Quotient (AQ-10), the Autism Spectrum Screening Questionnaire-Revised Extended Version (ASSQ-REV), the Developmental Behavior Checklist-Autism Screening Algorithm (DBC-ASA), the Diagnostic Behavioral Assessment for Autism Spectrum Disorder-Revised (DiBAS-R), A DSM-5 teacher screening questionnaire for autism & social communication disorders (EDUTEA), the Autism Diagnostic Inventory-Telephone Screening in Spanish (ADI-TSS), the Adapted Autism Behaviour Checklist (AABC) and the Mobile Autism Risk Assessment (MARA).

After evaluating the evidence for the twelve tools, two of them (SCQ and AQ-10) were selected for further review by a consensus group of Nigerian experts because the SCQ had evidence of cross-cultural use and the AQ-10 was the adolescent version. The group of experts examined the content and face validity of both tools. After the consensus group consultations in Nigeria (Nwokolo, Murphy, Mensink, Moonen & Langdon, submitted), the Social Communication Questionnaire (SCQ) was selected, and adjusted slightly to contain more culturally relevant examples. The SCQ is a brief 40-item parent or caregiver screening measure used widely in research (Berument, Rutter, Lord, Pickles & Bailey, 1999). Administration of the SCQ is time-efficient, requiring no costly or special training. The group agreed that the SCQ was more robust and comprehensive than the AQ-10, with questions that examined the relevant autism spectrum domains.

To establish the usefulness of the SCQ in Nigeria, this study aimed to a) validate the structure of the SCQ in the Nigerian population using confirmatory factor analysis (CFA), b) examine the

internal consistency, discriminative, and convergent validity of the SCQ, c) derive an appropriate cut-off score based upon sensitivity and specificity, in relation to the criterion measure (the ADOS-2) and d) derive the positive and negative predictive values.

The recruitment of participants from the relevant centres was purposive to allow the inclusion of some persons suspected to have autism spectrum disorder, and some thought not to have the disorder, given the study's aims.

Methods

Design

A between-groups design was then used with two groups of participants: adolescents and young people with and without suspected autism spectrum disorder.

Participants

An adolescent or young person was eligible to take part in this study in the 'suspected ASD' group, if they were a) between 11- and 26 years old, b) identified by a doctor as having a clinical diagnosis of ASD, and/or c) enrolled in a special education school or a special centre and d) had a parent, guardian, or caregiver with adequate lifetime information regarding the adolescent. The 'non-autistic' participants (i.e., those not suspected of having autism) were recruited. The study occurred within three of Nigeria's geopolitical zones: Abuja, Enugu, and Lagos. Participants were recruited from day centres, special schools, child and adolescent mental health care services, local community organisations, religious organisations, and public advertisements.

As a result, two hundred and ten adolescents and young adults, 124 (59%) male and 86 (41%) female (M_{age} = 15.88 years; Mdn_{age} = 15.29 years; SD = 3.69; range = 10.90 – 26.96 years) took part in this study. The age distribution was grouped as follows: 11 – 13-year-olds (n = 76; 36.2%), 14 – 15-year-olds (n = 42; 20%), 16 – 17-year-olds (n = 51; 24.3%) and 18 years and above (n = 41; 19.5%). Initially, 245 adolescents and young adults were invited to participate in this study, but 35 declined to participate or did not respond to further attempts to contact them, so that finally 210 took part. Out of the 210, a further 5 participants did not complete the SCQ and were excluded from the analysis leaving 205 as the total number of participants.

Procedure

Ethics

A positive ethical opinion was obtained from the University of Kent, Tizard Centre Ethics Committee, the National Health Research Ethics Committee (NHREC; NHREC/01/01/2007-16/09/2019) and the Federal Neuro-Psychiatric Hospital, Yaba, Lagos, Nigeria (FNPH/HREC/20/09). All participants were provided with written information about participating in this study, including easier-to-read versions of information sheets. Parental informed consent was sought for those aged under 18 years of age. Participants were encouraged to take breaks as needed during testing.

Setting and Procedure

Each participant was seen either in their school, place of worship, centre, clinic or home and was invited to complete the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2) while the SCQ was completed by their parent, caregiver, or guardian. Participants were included in the autism group if they met the autism or autism spectrum classification on the ADOS-2. The ADOS-2 was administered to the adolescent by the first author, who has been trained to meet the ADOS-2 to clinical reliability.

Measures

Social Communication Questionnaire (SCQ)

The SCQ is a brief 40-item parent or caregiver-report screening measure modelled after the Autism Diagnostic Interview-Revised (ADI-R) and has been used widely in research (Berument et al., 1999; Rutter et al., 2003). It is a screening tool only and cannot be used for the diagnosis of autism. The SCQ is designed for anyone 4 years old and above, and it takes about 10 – 15 minutes to complete and about 5 minutes to score. Following a consensus group study prior to this study, minor adjustments were made to examples in the SCQ, to ensure that it was culturally relevant (see Nwokolo et al, submitted). The measure has two versions: the lifetime and the current versions. Both focus on symptoms of autism most likely to be observed by the individual's principal caregiver, who must be familiar with the individual's developmental history and current behaviour. The lifetime version was used in this study, given the age range of the participants (11 – 26 years). In addition, Wei, Chestnut, Barnard-Brak & Richman (2015) reported that the lifetime version has better psychometric properties.

Autism Diagnostic Observation Schedule, Second Edition (ADOS-2)

The Autism Diagnostic Observation Schedule, 2nd Edition, is a very widely used tool for assessing the presence of autism (Chojnicka & Pisula, 2017; Lebersfeld, Swanson, Clesi & O'Kelley, 2021). It is a semi-structured standardised clinician tool which uses a hierarchy of presses across a range of play-based activities to observe behaviour, communication, social interaction and imaginative use of materials. An overall score is obtained with cut-offs for ASD. The ADOS-2 has five modules – toddler and modules 1 to 4. Modules 1, 2 and 3 were used in this study. To determine the applicable module for each participant, the suggested guidelines in the ADOS-2 manual was followed. The guideline includes evaluation of the individual's expressive language and determining the chronological age. The ADOS-2 takes between 60 – 90 minutes to administer and score.

According to NICE guidelines in the UK (NICE, 2017), the ADOS-2 should be used alongside the ADI-R or DISCO interview to make a certain diagnosis of autism. However, resources did not allow for the ADI-R or DISCO, so the ADOS-2 was used alone as the gold standard.

Data Analysis

Overview

The Statistical Package for the Social Sciences – IBM SPSS version 26 and Jeffreys's Amazing Statistics Program (JASP) version 0.16.3, an open-source statistical package, were used for analyses. Except for the confirmatory factor analysis (CFA) done with JASP, all other analyses were done with SPSS. CFA was performed to confirm the applicability and validity of the original SCQ constructs to the Nigerian adolescent population. The SCQ's performance as a screening tool was compared to the ADOS-2 classification, while correlations between the 40 SCQ items and ADOS-2 were calculated using Pearson's *r*. The discriminative and convergent validities of the SCQ were examined. Internal consistency was calculated using Cronbach's alpha. Receiver Operating Characteristic (ROC) analyses were used to calculate the Area under the Curve (AUC) to examine how well the SCQ identified participants with and without an autism spectrum disorder, with reference to sensitivity and specificity to identify optimal cut-offs. The positive (PPV) and negative predictive (NPV) values were calculated from the results.

Missing Data

5 participants did not complete the SCQ, and 6 did not complete the ADOS-2. Therefore 5 participants were excluded from the SCQ analysis and 6 from the ADOS-2 analysis.

Confirmatory Factor Analysis (CFA)

The SCQ is a standardised measure used widely in research. It has been translated into a number of different languages and used in a variety of countries, for example, in South Africa (Bozalek, 2013), Uganda (Awadu, 2021) and China (Liu et al., 2022). Some studies have examined the structure (Uljarević, Frazier, Phillips, Jo, Littlefield & Hardan, 2021), psychometric properties (Wei et al., 2015) and its utility as a screening tool (Chestnut, Wei, Barnard-Brak & Richman, 2016). Whereas some Nigerian professionals are conversant with the SCQ, the applicability of the SCQ has not been examined. Thus, a confirmatory factor analysis (CFA) was performed to evaluate the validity of the original SCQ constructs in the Nigerian adolescent population.

Initial model fit using the original four factors (social interaction, communication, abnormal language, and stereotypic behaviour, of Berument et al., 1999) was examined using the JASP Version 0.16.3. This was followed by a bootstrapping with replacement (based on a sample size of 5000) and the diagonal weighted least squares (DWLS) estimator, which are appropriate for small sample sizes (Mîndrilă, 2010; DiStefano & Morgan, 2014; Koğar & Koğar, 2015). The model fit was first evaluated based on the chi-square (χ^2) goodness-of-fit statistics. Due to the sample size, literature and various opinions and criticisms about the chi-square, other indices were also examined: the chi-square/*df* ratio <3, the comparative fit index (CFI) ≥ .90, the Tucker-Lewis index (TLI) ≥ .90, the root mean square error of approximation (RMSEA) ≤ .08, the goodness of fit index (GFI) ≥ .90 and the standardised root mean square residual (SRMR) between .05 and .08 (Hu & Bentler, 1999; Prudon, 2014; Newsom, 2018a, 2018b; Mîndrilă, 2010).

Internal Consistency

Cronbach's alpha, considered an adequate measure of internal consistency (Mokkink et al., 2018; Terwee et al., 2007), was used to assess the internal consistency of the SCQ. A Cronbach's alpha greater than or equal to 0.70 is acceptable (Tavakol & Dennick, 2011).

Criterion Validity

The criterion validity of the SCQ was determined by assessing both the discriminative and convergent validities. Terwee et al. (2007) and Mokkink et al. (2018) suggest a good correlation with the 'gold standard' tool if the correlation is ≥ 0.70 or AUC ≥ 0.70.

- Discriminative Validity. The discriminative ability of the SCQ was determined by examining the AUC.
- Convergent Validity**.** Convergent validity was assessed by examining the extent of correlation between the SCQ scores and the ADOS-2 classifications. The correlation was determined by using Pearson's correlation coefficient ***r***.

Sensitivity and Specificity

The sensitivity of the SCQ refers to the probability of it correctly identifying individuals with ASD, while the specificity refers to its probability of correctly identifying those who do not have ASD (Trevethan, 2017). The optimal cut-off score for the SCQ was based on the ROC analysis, while specificity and sensitivity were determined from the AUC (Streiner & Cairney, 2007; Lasko, Bhagwat, Zou & Ohno-Machado, 2005). Sensitivity and specificity cut-off values were guided by Glascoe (2005).

Positive Predictive Value (PPV) and Negative Predictive Value (NPV)

PPV and NPV determine those who genuinely have or do not have an ASD. The SCQ's PPV was determined using the formula [(true positives/(true positives + false positives)]*100 and NPV as [(true negatives/(true negatives + false negatives)]*100. Because both values relate to prevalence, there are no agreed cut-off values for PPV & NPV for a screening tool (Glascoe, 2005; Wong & Lim, 2011).

Results

Descriptive Statistics

SCQ scores for all 205 participants ranged from 0 to 30 points (M = 8.42; Mdn_{score} = 6.00; SD = 6.89). Total SCQ scores did not differentiate by sex (t_{205} = .34, p = .74). Descriptive statistics for the SCQ and ADOS-2 scores per age group are in Tables 1 and 2. The distribution of the 204 participants with and without ASD (as defined by the ADOS-2 autism spectrum cut off score for each module) is in Table 3.

Table 1 About Here

Table 2 About Here

Table 3 About Here

Confirmatory Factor Analysis (CFA)

Our CFA had an acceptable model fit, CFI = 1.00, TLI = 1.03, SRMR = 0.08, RMSEA = .00 and GFI = .96. Figure 1 shows the model plot and factor loadings in Table 4. The initial hypothesis was that the original four factors (social communication, social interaction, abnormal language, and stereotypic behaviour) of the SCQ may not be applicable in the Nigerian context, but it transpired that the original four factor structure could be maintained. Items 5, 9 and 13 had factor loadings that were slightly below 0.3 (0.26, 0.20 & 0.20, respectively) but were retained in the SCQ as removing was not deemed impactful to the structure of the SCQ. Overall, the factor loadings indicated that all the items could be retained in the SCQ.

Table 4 About Here

–

Figure 1 About Here

Internal Consistency

The Cronbach's α = .88 for the total sample and α = .86 for the ASD group on all original four domains of the SCQ indicated adequate internal consistency of the SCQ items, while for the non-ASD group α = .59. In the entire sample, Cronbach's α = .85 for the social communication and interaction domain and .66 for the restricted, repetitive and stereotypic pattern of behaviours domains (RRSB) – items 7, 8, 11, 12, 13, 14, 15 & 16. With the addition of the self-injurious items (17, 18 & 38) to the RRSB, Cronbach's α = .71.

Criterion Validity

Discriminative Validity

The cut-off score of 10 on the SCQ showed that the SCQ could differentiate between those with and without ASD, (using the cut-off score for autism spectrum applicable for each ADOS-2 module). An AUC of 1 would indicate a perfect screening tool. At the cut-off score of 10, the AUC was .83.

Convergent Validity

Overall convergent validity was indicated by a significant Pearson's correlation between the total SCQ scores and ADOS-2 scores for the 204 participants, $r = .71$, $p < 0.001$, showing a strong correlation and effect size (Cohen, 1992).

Sensitivity and Specificity

At the recommended cut-off score of 15, ROC analysis revealed an overall AUC for the 204 participants as .76, $p < 0.001$, 95% CI [.68, .84] with PPV = .54, NPV = .99, sensitivity = .95 and specificity = .81. While the specificity and sensitivity met the minimum requirements, the PPV did not meet the minimum standard (Glascoe, 2005). Lower cut-offs were explored by stepwise reduction to determine the best cut-off score. Lowering the cut-off score to 10, 11 and 12 improved the values; however, a cut-off score of 10 gave the best result overall, as shown in

Table 5. With a cut-off score of 10, sensitivity = .81 and specificity = .88, applicable to the entire population.

Table 5 About Here

For the specific age groups, using the cut-off score of 10, the following results were obtained: 11 – 13-year-olds, AUC = .83, $p < 0.001$, 95% CI [.77, .90], $N = 74$, 14 – 15-year-olds, AUC = .84, $p < 0.001$, 95% CI [.94, .99], $N = 40$, 16 – 17-year-olds, AUC = .84, $p < 0.001$, 95% CI [.71, .97], $N = 50$, 18 years and above, AUC = .83, $p < 0.001$, 95% CI [.67, 1.00], $N = 40$. Sensitivity and specificity with optimal cut-off scores for each group are shown in Table 4. Our study explored the usefulness of differentiated cut-off scores per age group since, to the best of our knowledge, no study specifying SCQ cut-off scores for age brackets has been done. From our study, a cut-off score of 10 is preferable for all those under 18yrs, and a cut-off score of 12 is best suited for participants aged 18 and above, as all the psychometric properties met the minimum standard (Table 5).

Positive Predictive Value (PPV) and Negative Predictive Value (NPV)

Using the SCQ cut-off score of 10, PPV = .75 and NPV = .91, showing that the SCQ can correctly identify those with ASD and those without ASD.

Discussion and Conclusion

With the increased awareness of autism spectrum disorder in Nigeria, parents of younger children now seek screening and early intervention. However, the older children and adolescents who missed early screening and diagnosis need to be known. To detect ASD in these adolescents, a validated and easy-to-use screening tool is required. The SCQ was identified through a systematic review (Nwokolo et al., 2022) and agreed on as appropriate by a focus group (Nwokolo et al., submitted). Thus, the goals of this study were to a) validate the structure of the SCQ in the Nigerian population using confirmatory factor analysis (CFA), b) examine the internal consistency, discriminative, and convergent validity of the SCQ, c) derive an appropriate

cut-off score based upon sensitivity and specificity and d) derive the positive and negative predictive values.

Confirmatory Factor Analysis (CFA)

While the SCQ is an established and widely used measure in both research and clinical settings, the accuracy and psychometrics of the SCQ have mainly been examined in North and South America, Europe, and Australia (Chestnut et al., 2017). Studies confirming its appropriateness in the African context, especially among adolescents, were non-existent. However, studies which examined the discriminative validity in young and older children aged between 2.5 and 17 years in South Africa (Bozalek, 2013) and Uganda (Awadu, 2021) were found. Based on existing literature, the scarcity of cross-cultural research in the African context and the aims of this study, a CFA was done. The CFA results revealed that the Nigerian population could retain the original four-factor structure, bearing in mind the limitations of the sample size.

Internal Consistency

The SCQ's internal consistency was adequate, with a Cronbach's alpha of .88, indicating the tool's ability to capture the concept of autism spectrum disorder.

Criterion Validity

Although at the published cut-off score of 15, the sensitivity, specificity and NPV met recommended criteria (Glascoe, 2005), the PPV was not optimal. With a reduction in the cut-off score to 10, all the properties met the minimum criteria, with the SCQ adequately discriminating between those with ASD and those without ASD. There was a strong positive relationship between the SCQ scores and group classification (with or without ASD), which showed that as the SCQ scores increased, the more likely an individual would have an ASD. The SCQ, as a screening tool, correlated well with the ADOS-2 (r = .70), showing that the SCQ is a valid screening instrument for use with the Nigerian adolescent population.

Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value

Based on the results of the ROC analyses, the sensitivity, specificity, PPV and NPV met the acceptable criteria of 70% for screening tools (Glascoe, 2005). Initially, at the recommended cut-off score of 15, PPV = .54, NPV = .99, sensitivity = .95 and specificity = .81. However, the discriminating ability improved by reducing the cut-off to 10, giving a sensitivity of .81 and

specificity of .88. This cut-off score of 10 is like the results obtained by Bozalek (2013) in the South African sample (cut-off = 10, sensitivity = 1, specificity = .95) and Awadu, 2021) in Uganda (cut-off = 10, sensitivity = 1, specificity = .93). Other studies (Kim et al., 2015; Snow & Lecavalier, 2008; Schanding, Nowell & Goin-Kochel, 2012) also recommended a reduction in the cut-off score from 15 for better outcomes. In assigning the cut-off, the distinctions between sensitivity and PPV and between specificity and NPV in a screening and clinical context were considered (Trevethan, 2017; Akobeng, 2007). Classifying participants solely on sensitivity and specificity values differs from classifying them in combination with the PPV and NPV. PPV and NPV are influenced by the prevalence and depend on the population being investigated. Participants identified by medical professionals as autistic were sampled in this study; as such, the PPV and NPV were considered in addition to sensitivity and specificity to determine the best cut-off score. Since the Lifetime version of the SCQ was used, it is possible that some of the respondents of the older participants may not have an absolute recollection of the early years of their wards. For this reason, the scores will be affected, and a lower cut-off ensures that persons who may have autism are not missed. Should the sample age in any study be homogenous, which is highly unlikely, specific cut-offs are recommended for the different age groups, as shown in Table 6. Overall, the results showed that the SCQ correctly identified adolescents with and without ASD.

Table 6 About Here

Limitations

There are some limitations to this study. First, the sample size was relatively small for CFA, although our model was associated with a good fit. Secondly, the participants were mainly from urban settings and had good literacy skills; thus, it cannot be assumed that the psychometric properties will be the same when used in rural settings, where questions may need to be read to respondents. Thirdly, while we recognise that the use of the original English version of the SCQ may be judged insufficient for a validation study in an ethnic and culturally diverse setting as Nigeria, English is the official language in Nigeria, and we ensured the examples given were culturally appropriate. English as the official language, or pidgin (a variation of English) is

widely spoken by most people especially the urban dwellers. Additionally, while urban populations may be similar, given that there is insufficient evidence of formal validation of any autism screening tool for the Nigerian population, validation of English SCQ was deemed a viable start. Further studies to explore the translation and validation of the SCQ in the three major languages (Hausa, Igbo and Yoruba) is recommended. Fourthly, we categorised participants as autistic based upon the ADOS-2 only and we did not undertake an additional assessment such as the ADI-R, DISCO or generate a thorough developmental history. A similar criticism about the use of English may arise concerning the use of the ADOS-2, however, for the same reasons that the English SCQ was used, and in the absence of other available tools, the ADOS-2 was deemed appropriate for use with the Nigerian urban population. It is possible that in doing so, functional and stimulus biases may have been introduced in that the Nigerian participants may not have been offered the same opportunity to demonstrate knowledge while eliciting the intended response as participants in the original ADOS-2 study. A study examining the validity of the ADOS-2 in the Nigerian context is recommended. Despite these limitations, the SCQ appears to be a useful screening tool for ASD in Nigerian adolescents.

Conclusion

In conclusion, the SCQ Lifetime form's psychometric properties met acceptable screening tools standards across the entire sample and all age groups of Nigerian adolescents and young people. All items of the SCQ Lifetime version are relevant, with culturally relevant examples used as applicable. Based on available data, this study is the first to explore the usefulness of differentiated cut-off scores per age group for the SCQ. From this study, a cut-off score of 10 is recommended for all those under 18yrs, and a cut-off score of 12 for participants aged 18 and above, as all the psychometric properties met the minimum standard. Further studies exploring these cut-offs are recommended. The SCQ Lifetime form can be used as a screening tool for identifying Nigerian adolescents likely to have autism spectrum disorder and help ensure referrals for further diagnosis. Using the suggested cut-offs for specific age groups will be beneficial in clinical settings.

Financial Disclosure

None

Funding Source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors do not have any conflict of interest to declare.

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*.

Awadu, J. E. (2021). 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. Michigan State University.

Bakare, M. O., Frazier, T. W., Karpur, A., Abubakar, A., Nyongesa, M. K., Mwangi, P. M., Dixon, P., Khaliq, I., Gase, N. K., Sandstrom, J., Okidegbe, N., Rosanoff, M., Munir, K. M., Shih, A. (2022). Brief report: Validity and reliability of the Nigerian autism screening questionnaire. *Autism*, 13623613221080250.

Bello-Mojeed, M., Bakare, M., & Munir, K. (2013). Identification of autism spectrum disorders (ASD) in Africa: Need for shifting research and public health focus. *The Comprehensive Guide to Autism, 12*, 43.

Bello-Mojeed, M., Omigbodun, O., Bakare, M., & Adewuya, A. (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.*

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism Screening Questionnaire: Diagnostic validity. *British Journal of Psychiatry, 175*(5), 444-451.

Bozalek, F. (2013). Autism screening in children: using the Social Communication Questionnaire in a Western Cape population. (Thesis). University of Cape Town, Faculty of Humanities, Department of Psychology.

Burkett, K., Morris, E., Manning-Courtney, P., Anthony, J., & Shambley-Ebron, D. (2015). African American families on autism diagnosis and treatment: The influence of culture. *Journal of Autism and Developmental Disorders, 45*(10), 3244-3254.

Chojnicka, I., & Pisula, E. (2017). Adaptation and validation of the ADOS-2, Polish version. *Frontiers in Psychology, 8*, 1916.

Delehanty, A., Lee, J., Hooker, J. L., Cortese, J., & Woods, J. (2020). Exploring message framing to engage parents in early screening for autism spectrum disorder. *Patient Education and Counseling, 103*(12), 2525-2531.

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(3), 425-438.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53-60.

Horlin, C., Falkmer, M., Parsons, R., Albrecht, M. A., & Falkmer, T. (2014). The cost of autism spectrum disorders. *PloS One, 9*(9), e106552.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

James, S. N., & Smith, C. J. (2020). Early autism diagnosis in the primary care setting. Paper presented at the *Seminars in Pediatric Neurology, 35* 100827.

Kim, J., Sunwoo, H., Park, S., Noh, D., Jung, Y. K., Cho, I., Cho, S., Kim, B., Shin, M., Kim, J., Park, T., Son, J., Chung, U., & Yoo, H. J. (2015). A validation study of the Korean version of Social Communication Questionnaire. *Journal of the Korean Academy of Child and Adolescent Psychiatry, 26*(3), 197-208.

Koğar, H., & Koğar, E. Y. (2015). Comparison of different estimation methods for categorical and ordinal data in confirmatory factor analysis. *Journal of Measurement and Evaluation in Education and Psychology, 6*(2)

Lebersfeld, J. B., Swanson, M., Clesi, C. D., & O'Kelley, S. E. (2021). Systematic review and meta-analysis of the clinical utility of the ADOS-2 and the ADI-R in diagnosing autism spectrum disorders in children. *Journal of Autism and Developmental Disorders*, *51*(11), 4101-4114.

Liu, S., Wang, X., Chen, Q., Chen, J., Jin, C., Zhan, X., Guo, C., Li, X., Lin, L., & Jing, J. (2022). The validity and reliability of the simplified Chinese version of the Social Communication Questionnaire. *Autism Research*, 1-10.

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). ADOS-2: Autism Diagnostic Observation Schedule, (ADOS-2). *Torrance, CA: Western Psychological Services.*

McCarty, P., & Frye, R. E. (2020). Early detection and diagnosis of autism spectrum disorder: Why is it so difficult? Paper presented at the *Seminars in Pediatric Neurology*, *35* 100831.

Mîndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, *1*(1), 60-66.

Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2018). COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). *User manual. https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf* Accessed 28 November 2018.

Muthén, B. (1993). Goodness of fit with categorical and other nonnormal variables. S. 205-234 in: Bollen, KA and long, JS (Hrsg.), testing structural equation models.

Nadel, S., & Poss, J. E. (2007). Early detection of autism spectrum disorders: Screening between 12 and 24 months of age. *Journal of the American Academy of Nurse Practitioners*, *19*(8), 408-417.

Newsom, J. (2018). Alternative estimation methods (psy 523/623 structural equation modeling, spring 2018). *Manuscript Retrieved from Http://web.Pdx.edu/~newsomj/semclass/*,

Newsom, J. (2018). Some clarifications and recommendations on fit indices (psy 523/623 structural equation modeling, spring 2018). *Manuscript Retrieved from Http://web.Pdx.edu/~newsomj/semclass/*,

National Institute for Health and Care Excellence. (2017). "*Autism spectrum disorder in under 19s: recognition, referral and diagnosis.*" (CG128). ttps://www.nice.org.uk/guidance/cg128 Accessed 11 February 2023.

Nwokolo, E. U., Langdon, P. E., & Murphy, G. H. (2022). Screening for intellectual disabilities and/or autism amongst older children and young adults: A systematic review of tools for use in Africa. *Review Journal of Autism and Developmental Disorders*. Advanced online publication.

Nwokolo, E. U., Murphy, G. H., Mensink, A. & Moonen, X. M. H., Langdon, P. E., (submitted). Using the consensus group method to select the best screening tools for autism and intellectual disability for use with Nigerian adolescents. *Journal of Policy & Practice in Intellectual Disabilities.*

Oshodi, Y. O., Olagunju, A. T., Oyelohunnu, M. A., Campbell, E. A., Umeh, C. S., Aina, O. F., Adeyemi, J. D. (2017). Autism spectrum disorder in a community-based sample with neurodevelopmental problems in Lagos, Nigeria. *Journal of Public Health in Africa, 7*(2), 559; 559-559.

Penner, M., Rayar, M., Bashir, N., Roberts, S. W., Hancock-Howard, R. L., & Coyte, P. C. (2015). Cost-effectiveness analysis comparing pre-diagnosis autism spectrum disorder (ASD)-targeted intervention with Ontario's autism intervention program. *Journal of Autism and Developmental Disorders, 45*(9), 2833-2847.

Prudon, P. (2014). Confirmatory factor analysis: A brief introduction and critique. *Qualtrics, P.UT, USA.*

Rosenberg, R. E., Landa, R., Law, J. K., Stuart, E. A., & Law, P. A. (2011). Factors affecting age at initial autism spectrum disorder diagnosis in a national survey. *Autism Research and Treatment, 2011:*1–11.

Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism Diagnostic Interview-Revised. *Los Angeles, CA: Western Psychological Services, 29*(2003), 30.

Sampaio, F., Feldman, I., Lavelle, T. A., & Skokauskas, N. (2021). The cost-effectiveness of treatments for attention deficit-hyperactivity disorder and autism spectrum disorder in children and adolescents: A systematic review. *European Child & Adolescent Psychiatry*, 1-16.

Schanding, G. T., Nowell, K. P., & Goin-Kochel, R. P. (2012). Utility of the Social Communication Questionnaire-current and Social Responsiveness Scale as teacher-report screening tools for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(8), 1705-1716.

Shevlin, M., & Miles, J. N. (1998). Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual Differences, 25*(1), 85-90.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34-42.

Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health, 5*, 307.

Wei, T., Chesnut, S. R., Barnard-Brak, L., & Richman, D. (2015). Psychometric analysis of the Social Communication Questionnaire using an item-response theory framework: Implications for the use of the lifetime and current forms. *Journal of Psychopathology and Behavioral Assessment, 37*(3), 469-480.

Wing, L., & Potter, D. (2002). The epidemiology of autistic spectrum disorders: Is the prevalence rising? *Mental Retardation and Developmental Disabilities Research Reviews, 8*(3), 151-161.

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A. & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research, 15*(5), 778-790.

Appendix 24 - SCIL

Screener for Intelligence and Learning Disabilities (SCIL)

1. Do you receive special education? Do you go to a special needs school? Do you have a special educational needs (SEN)?

2. Which school/college do you attend now, or did you attend in the past?
    a. None
    b. Primary school
    c. Special needs school
    d. WAEC/IGCSE/SAT
    e. A-level
    f. Polytechnic/Monotechnic/Teacher's college
    g. University
    h. Other

3. Do you receive or have you received support from a service for people with Intellectual Disability (excluding a home tutor or lesson teacher)?

4. Have you got family members, relatives or friends who you can contact if you have a problem (for example a difficult situation or emergency)?

5. In the supermarket you need to pay 6.95. How much change would you get back from 10?

6. Imagine you are at your Doctors on the 19th of January. He or she wants to see you again in three weeks. When (which date) would that be?

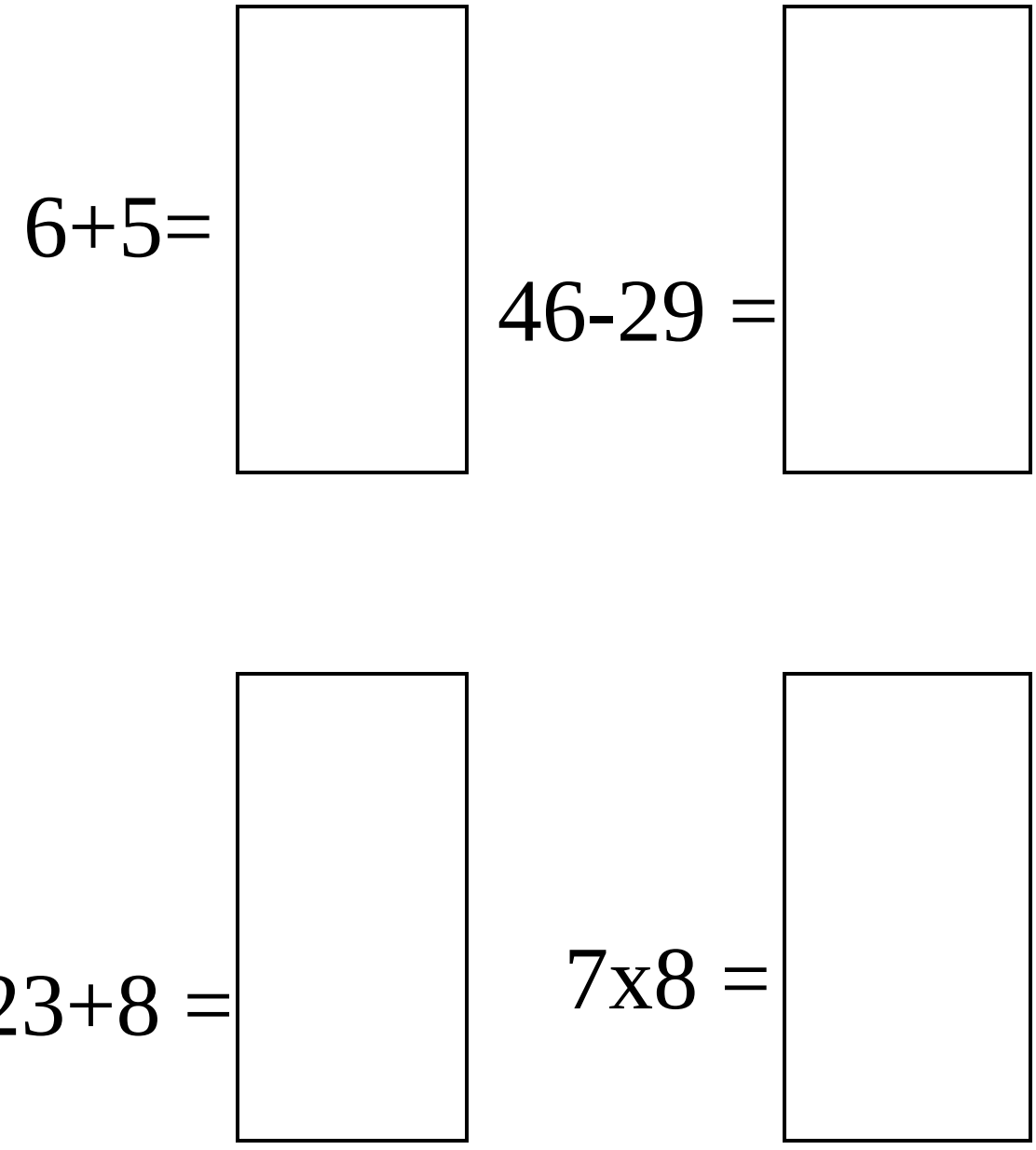
7. Imagine you are at your Doctors on the 3rd of January. He or she wants to see you again in three weeks. When (which date) would that be?

8. Can you say this word backwards? Say every letter STORM

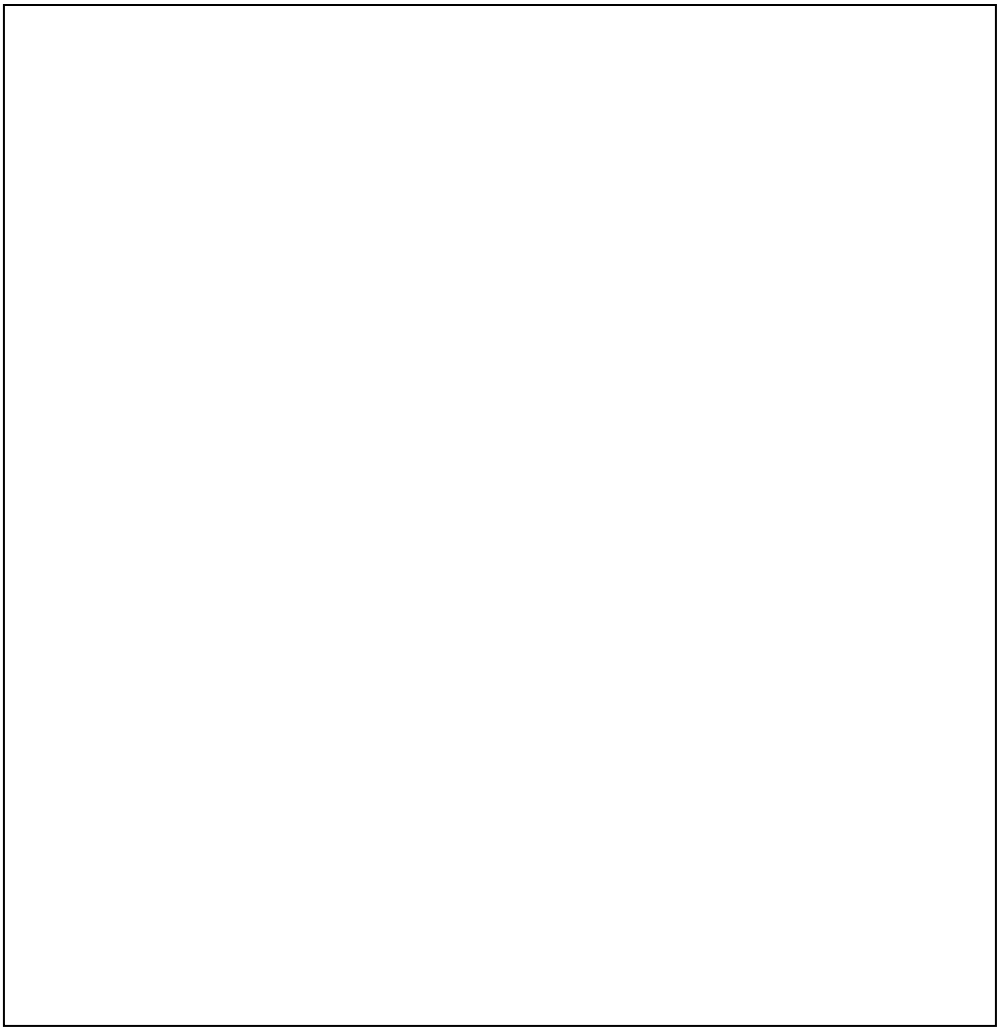
9. Do you read a newspaper or magazine? If so, which one?

10. What does this mean: "Like father, like son?"

11. Arithmetic – On this sheet are four sums. You need to write down the answer for every sum.



12. Writing – I'm going to read a few sentences for you to write in the box. Try to do this well or correct and as fast as you can.
    a. We are dumping a load of sand in the back garden.
    b. During the night the driver had to avoid knocking down a cow.



13. Reading – I'm going to ask you to read a story. Read this as quickly as you can without making mistakes.

    It is possible to pay for parking with your bank card.

    When you have parked your car, you use your bank card to pay as

    advertised/displayed on the signs and parking machines. When you leave you take your receipt.

14. Clock drawing – In this box draw a clock that says 9:45 (quarter to ten). Draw this as complete/detailed as you can with hands.