Evaluating the Psychometric Properties of Neurodevelopmental Assessment Tools in a South African Context

Michal R. Zieff

ZFFMIC001

Dissertation presented for the degree of Master of Science (Medicine) in the Division of Developmental Paediatrics, Department of Paediatrics and Child Health, Faculty of Health Sciences, University of Cape Town

August 2021

Supervisors

Professor Kirsten A. Donald

Dr Michelle Hoogenhout

Declaration

I, Michal R. Zieff, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents.

Signature: Signed by candidate

Date: 2 August 2021

Contents

List of Tables

Chapter 2

Chapter 4

List of Figures

Chapter 2

Chapter 3

Chapter 4

Acknowledgements

I acknowledge and thank all the children and families who generously gave of their time to participate in the study and taught me so much in the process.

It has been a great privilege to work with and learn from my supervisor, Professor Kirsty Donald. I thank Kirsty for her outstanding knowledge contributions to this dissertation and for her continued support. I am also very grateful to my co-supervisor, Dr Michelle Hoogenhout, for her support and invaluable mentorship over the past few years. This project has also benefited greatly from Michelle's knowledge in the field of neuropsychology and her expertise in data science.

I could not have completed this degree without the generous fellowship funding from the Harry Crossley Foundation. This research was also funded by grant awards received from the University of Cape Town's Faculty of Health Sciences Research Committee and Department of Paediatrics and Child Health Research Committee. Thank you for believing in me and the project.

I am so grateful to the entire NeuroDEV South Africa data collection team – Björn, Claire, Emma, Fatima, Gina, Jess, Lizette, Lopo, Megan, Nonceba, Raphi, and Serini – for their hard work and support, and to the wonderful Greta from S6 Developmental Clinic at Red Cross Children's Hospital for her patient assistance. A special thanks to Emma Eastman for her tireless efforts to ensure the smooth running of the study and for being an all-round saviour. I could not have reached this point without Emma's patience and generosity. I also thank the NeuroDEV study team from the Stanley Centre for Psychiatric Research at the Broad Institute of MIT and Harvard, and the KEMRI-Wellcome Trust Collaborative Research Programme in Kilifi, Kenya, especially Drs Elise Robinson, Victoria de Menil, Charles Newton, and Amina Abubakar for their generous intellectual contributions to my work.

I also acknowledge and thank the following people for their contributions to the project: Professor Colin Tredoux (Department of Psychology, University of Cape Town), for sharing insights on the statistical aspects of the study; Robbie Katz, for creating Python code that made data wrangling more efficient; Mrs. Mary Shelton (University of Cape Town Health Sciences Library), for invaluable guidance with the systematic review study.

Finally, I thank CSM and my family for their unconditional love, as well as the support team – Jenny, Emma, and Kathy – for their care, reassurance, and encouragement over the years.

Abstract

Neurodevelopmental and behavioural tools, used to assess children's cognitive and psychological development, play important supportive roles in clinical decision-making processes. It is therefore important that inferences made based on the information generated by these tools are valid and reliable. Because validity and reliability are inextricably tied to the context in which the tool is administered, it is critical that tools, including established "gold standard" tools, undergo ongoing psychometric evaluation. This is particularly important when tools are used in contexts different to that in which the original tool was developed. The use of poorly performing measures in clinical settings may result in a higher risk of misinterpretation of results or misdiagnosis, carrying serious implications for provision of interventions.

This dissertation explored the psychometric properties of two frequently used behavioural tools in a South African context. The first aim of the dissertation was to investigate the use, cultural appropriateness, and psychometric soundness of Achenbach System of Empirically Based Assessment (ASEBA) forms in sub-Saharan Africa. The ASEBA forms are used worldwide to screen children and adolescents for behavioural and emotional problems. To achieve this aim, I conducted a systematic review of the psychometric properties of the ASEBA forms in sub-Saharan Africa. The second aim was to evaluate the structural and construct validity of a commonly used measure of Attention-Deficit/Hyperactivity Disorder (ADHD), the Swanson, Nolan, and Pelham ADHD Rating Scale (SNAP-IV), in a sample of South African children with neurodevelopmental disorders (NDDs), including Autism Spectrum Disorder and Intellectual Disability.

The systematic review identified 58 studies with sub-Saharan African participants that reported measurement properties of the ASEBA forms. Most studies came from Southern (n = 29, 50%) or East African (n = 25, 43%) countries. Forty-nine studies (84%) used translated versions of the tool, but details regarding the translation process, if available, were often sparse. Most studies (n = 47, 81%) only reported internal consistency (using coefficient alpha) for one or more subscales. The methodological quality of the psychometric evaluations varied considerably across all measurement properties, except for internal consistency. There is limited good quality psychometric evidence available for the ASEBA forms in sub-Saharan Africa. Recommendations include implementing a standardised procedure for conducting and reporting translation processes and conducting more comprehensive psychometric evaluations of the translated versions of the tools.

Parents of 109 children with one or more diagnosed NDDs completed the SNAP-IV. A subset of parents (n = 79) also completed the ASEBA Child Behaviour Checklist (CBCL/6-18). We conducted a confirmatory factor analysis to inspect the two-factor structure of the SNAP-IV (Inattention + Hyperactivity-Impulsivity). We also calculated ordinal coefficient alpha (α) to estimate internal consistency. Finally, we correlated scores of SNAP-IV and CBCL/6-18 subscales to estimate concurrent, convergent (ADHD + externalizing behaviour), and discriminant (ADHD + internalizing behaviour) validity respectively. The two-factor model performed acceptably (χ^2 (134) = 249.82, $p < 0.001$, TLI = 0.868, RMSEA = 0.089, $p < 0.001$). The model fit improved after removing three problematic items, two of which were dependent on the child's level of spoken language (χ^2 (89) = 135.17, $p < 0.01$, TLI = 0.942, RMSEA = 0.069, $p = 0.096$). The revised SNAP-IV subscales had acceptable internal consistencies (α = 0.85-0.86). Correlation coefficients between the SNAP-IV and ADHD-related CBCL/6-18 subscales were significant (r = 0.53-0.62, $p < 0.001$). Correlations between ADHD and externalizing behaviours (r = 0.45, $p < 0.001$) and internalizing behaviours (r = 0.38, $p < 0.001$) respectively were not significantly different (z = 0.97, p = 0.165). The findings tentatively support the use of the SNAP-IV in this group of children. However, there are limitations to its performance in this population likely related to the presence of NDDs.

Taken together, the findings of these two studies highlight the need for clinicians and researchers to conduct ongoing psychometric testing of behavioural tools for use with linguistically and culturally diverse sub-Saharan African populations. The data also reveal important insights regarding problems associated with using standard behavioural tools in children with complex clinical presentations.

Abbreviations

ADHD	Attention-Deficit/Hyperactivity Disorder
ASD	Autism Spectrum Disorder
ASEBA	Achenbach System of Empirically Based Assessment
CBCL	Child Behaviour Checklist
CD	Communication Disorder
CFA	Confirmatory factor analysis
COSMIN	Consensus-based Standards for the Selection of Health Measurement Instruments
DSM-5	Diagnostic and Statistical Manual of Mental Disorders (5th edition)
GDD	Global Developmental Delay
ID	Intellectual Disability
NDD	Neurodevelopmental disorder
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SLD	Specific Learning Disorder
SNAP-IV	Swanson, Nolan, and Pelham ADHD Rating Scale (4th edition)
SSA	sub-Saharan Africa
YSR	Youth Self-Report

Style

This dissertation includes two manuscripts that were submitted to an American and British journal respectively. To maintain consistency, I have used British spelling throughout the dissertation. I have also used the American Psychological Association (APA) 7th edition style for referencing and general formatting.

Glossary of Terms

I briefly define terms that are frequently used in this dissertation. Most of these terms are further explicated in Chapters 1-4. I divide them into three broad categories: Neurodevelopmental terms, psychometric terms, and statistical terms.

Neurodevelopment terms

Neurodevelopmental disorders (NDDs) are a group of conditions with onset in the developmental period (i.e., in early childhood), characterized by developmental delays or deficits in cognitive, social, or adaptive functioning. The following terms describe some of the NDDs listed in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) and other relevant concepts (American Psychiatric Association, 2013).

Attention-Deficit/Hyperactivity Disorder	An NDD characterised by impairments in attention and/or hyperactivity-impulsivity.
Autism Spectrum Disorder	An NDD characterised by deficits in social communication as well as repetitive behaviours, restricted interests, and insistence on sameness.
Cognitive age	An individual's level of cognitive functioning compared to the average cognitive functioning corresponding to the individual's chronological age (Kaplan & Saccuzzo, 2018).
Communication Disorders	NDDs characterised by deficits in the development and use of language, speech, and communication.
Global Developmental Delay	An NDD characterised by failure to meet expected developmental milestones in several areas of cognitive functioning (e.g., language, fine motor). This diagnosis is generally reserved for individuals who are unable to undergo standardised assessments of cognitive functioning (e.g., due to young age).
Intellectual Disability	An NDD characterised by deficits in general intellectual ability that result in impairments of adaptive functioning.

Specific Learning Disorder	An NDD characterised by specific deficits in learning and/or using one or more academic skills (e.g., reading).

Psychometric terms

Psychometrics is the science of psychological measurement, in particular the creation, application, and evaluation of assessment tools (Rust et al., 2021). The following terms describe a number of measurement properties (characteristics of scores derived from a psychological tool) and other important psychometric-related concepts. Mokkink and colleagues (2010) have published an excellent summary of the different domains and characteristics of important measurement properties (e.g., validity and reliability).

Construct bias	Construct bias may be introduced when a construct being measured by a tool is not understood or experienced in the same way across different cultural groups (Van De Vijver & Leung, 2011).
Cross-cultural validity	The degree to which the performance of a translated or adapted tool adequately reflects that of the 'original' tool in its 'original' context (Mokkink et al., 2010).
Internal consistency	Internal consistency refers to how well items within a test correlate with each other. If a test measures a single construct, and all items that make up the test are equally good at measuring this construct, then the test is internally consistent (R.B. Kline, 2020).
Latent trait	A theoretical construct that is not directly observable.
Likert scale	An item response format for which respondents indicate their level of agreement or disagreement with a statement, using an ordinal/ranked scale (e.g., 1 = strongly disagree; 5 = strongly agree).
Measurement bias	Any source of construct-irrelevant variance (i.e., variance not due to the construct being measured) that results in systematic error or inaccuracy in measurement (De Kock et al., 2013).

Normative data	Normative data refers to performance on a test by a defined group of people. The norms for a test (i.e., what is 'normal' or 'typical') are based on the distribution of the group's performance. A test score of an individual from the same defined group can then be interpreted relative to the performance of the normative sample (Kaplan & Saccuzzo, 2018).
Ordinal scale	A measurement scale that describes ranked variables, that is, variables with a natural order. Ordered variables have no true zero (i.e., no starting point) and therefore differences between the ranks are not mathematically meaningful.
Reliability	The extent to which a tool is free from measurement error or bias (Mokkink et al., 2010).
Inter-rater reliability	The degree of agreement between two (or more) independent examiners who rate or score the same behaviour. Inter-rater reliability is calculated by correlating the two sets of scores/ratings of the two examiners (Rust et al., 2021).
Test-retest reliability	The degree of agreement (i.e., correlation) between two sets of scores from the same tool that was administered to the same respondents on two separate occasions (Rust et al., 2021).
Validity	The extent to which a test measures the underlying theoretical construct that it purports to measure (Kaplan & Saccuzzo, 2018).
Concurrent validity	A type of construct validity describing the correlation between two measures intended to measure the same construct (Rust et al., 2021). We expect correlations between two tools purported to measure the same construct to be large and positive.

Content validity	The degree to which the content of a psychological tool is an adequate reflection of the construct being measured (Mokkink et al., 2010)
Convergent validity	A type of construct validity describing correlations between two constructs that should be related based on theory. We expect correlations between related variables to be positive and sizable (R.B. Kline, 2016).
Criterion validity	The extent to which the scores on a psychological tool are an adequate reflection of a "gold standard" criterion measure (Mokkink et al., 2010).
Discriminant (divergent) validity	A type of construct validity describing correlations between two constructs that are unrelated. We expect correlations between unrelated variables to be small, or at least smaller than those between related variables (R.B. Kline, 2016).
Known-groups validity	A type of construct validity concerned with a tool's ability to distinguish between groups known to vary on the construct being measured (de Vet et al., 2011).
Structural validity	The extent to which items on a tool can be used to measure the underlying theoretical construct. Structural validity is frequently explored using factor analytic techniques (De Kock et al., 2013). See 'Confirmatory factor analysis' and 'Exploratory factor analysis' on page xii.
Standardisation	Standardisation of a test occurs when a large number of individuals, from a variety of defined groups, are included in the normative sample to produce stratified normative data (P. Kline, 2000).

Statistical terms

The terms below describe core statistical techniques frequently used in psychometric evaluations.

Confirmatory factor analysis (CFA)	CFA is a statistical modelling technique used to examine relationships between observed variables (tool items) and latent factors (hypothetical constructs), based on an a priori specification about the underlying structure of a tool (R.B. Kline, 2016).
Cronbach's alpha	Cronbach's alpha, also known as coefficient alpha, is a statistical estimate of internal consistency reliability (Cortina, 1993). See 'Internal consistency' on page ix.
Exploratory factor analysis (EFA)	Like CFA, EFA estimates the relationships between observed and latent variables. However, EFA modelling is free to estimate any parameter (e.g., relationship between latent and ordinal variables) as it does not rely on an a priori specification (R.B. Kline, 2016).
Inter-item correlation	Correlation between the scores of two items on a scale.
Item-total correlation	The correlation between the scores on a test item and the total test score.
Ordinal coefficient alpha	An estimate of internal consistency reliability that takes into account the ordinal nature of Likert response data (Zumbo et al., 2007).
Polychoric correlation	A polychoric correlation estimates the linear relationship between two observed ordinal variables, each assumed to measure an underlying continuous latent variable (R.B. Kline, 2016)
Receiver Operating Characteristic (ROC)	ROC calculates the degree to which a tool discriminates between individuals based on a binary "gold standard" criterion (de Vet et al., 2011).

References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Author.

Cortina, J. M. (1993). What Is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

De Kock, F., Kanjee, A., & Foxcroft, C. (2013). Cross-cultural test adaptation, translation and tests in multiple languages. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed., pp. 102–125). Oxford University Press.

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge University Press.

Kaplan, R. M., & Saccuzzo, D. P. (2018). *Psychological testing: Principles, applications, & issues* (4th ed.). Cengage Learning.

Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). Routledge.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Kline, R. B. (2020). *Becoming a behavioral science researcher: A guide to producing research that matters* (2nd ed.). The Guilford Press.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, *63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Rust, J., Kosinski, M., & Stillwell, D. (2021). *Modern psychometrics: The science of psychological assessment* (4th ed.). Routledge.

Van De Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17–45). Cambridge University Press.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. https://doi.org/https://doi.org/10.22237/jmasm/1177992180

CHAPTER 1

Introduction

Parent-report neurodevelopmental and behavioural tools are used worldwide to assess children's cognitive and psychological development (Frick et al., 2020). Such tests play an important role in various decision-making processes. For example, information gathered from neurodevelopmental assessments help clinicians make diagnoses of neurodevelopmental disorders (NDDs) and decide on an appropriate course of therapeutic support for a child as needed, including intervention programmes and suitable schooling arrangements. Hence, it is of great importance that the measures used to assess children's development and behaviour are valid and reliable (Hunsley & Mash, 2007).

Many widely used neurodevelopmental tools are developed in the global North and thus are designed, standardised, and normed for individuals who predominantly come from urbanised, Westernised, English-language backgrounds and who are highly educated and literate (Fernald et al., 2009; Nezafat Maldonado et al., 2019; Sharp et al., 2014). It follows, then, that these tools may not produce reliable results when used with individuals who come from different sociocultural and language backgrounds (Sweetland et al., 2014). This is the case in South Africa, where English is spoken as a first language by only 13.5% of the population and where quality of education and degree of urbanisation vary considerably within the country (Nell, 2000; Shuttleworth-Edwards et al., 2004; Statistics South Africa, 2012). In addition, many South African individuals subscribe to systems of values and knowledge that are different to those typical of individuals living in the global North, which may also influence the validity and reliability of psychological tools (Grieve & Foxcroft, 2013). Therefore, the use of such tools in multicultural and multilingual South Africa may be problematic. Two negative consequences of using culturally or linguistically biased measures and inappropriate population norms are (i) the risk of misdiagnosis, which carries serious implications for the provision of treatment, interventions, and care dependency grants, and (ii) drawing invalid conclusions about populations based on the outcomes of tools in research studies. These consequences are especially concerning in resource-limited countries, such as South Africa, where funds and services are scarce.

In South Africa, tools are frequently translated into local languages for clinical and research purposes (De Kock et al., 2013). There are, however, several challenges associated with translating a tool into another language which may, in turn, compromise its validity and reliability (Van Widenfelt et al., 2005). For example, a lack of equivalent terms in the target language may complicate translation. Van Eeden and Mantsha (2007), in their study of a

Venda translation of the Sixteen Personality Factor Questionnaire (16PF), found that a lack of equivalent or similar words in Venda to describe various psychological constructs (e.g., depression, 'happy-go-lucky') may have altered the meanings of test items, and therefore, responses to those items. Ambiguous and multiple meanings of words in a target African language may also result in misinterpretation of test items. In addition, words in a target African language may be obscure or not of the same level of difficulty as the equivalent English word. One small study found that at least 10 out of the 42 word items on a Xhosa translation of a South African-adapted Wechsler Abbreviated Scale of Intelligence Vocabulary subtest had alternate meanings in addition to the intended meanings that were not accounted for in the scoring rubric (Zieff, 2017). Another similar study found that the item difficulty sequence on an Afrikaans translation of the Wechsler Adult Intelligence Scale-III Vocabulary subtest deviated from the expected pattern reflecting progression from easier to more difficult items (Grieve & Van Eeden, 2010). A group of local researchers translated the Western Aphasia Battery into Zulu, and found that the Zulu-translated items were occasionally more complex and prone to ambiguity than the original English items (Barratt et al., 2012). Overall, then, when directly translating an English test, failure to consider semantic and grammatical nuances of the target language may compromise a translated test's validity and reliability.

Problems with using tools developed in the global North in a multicultural context are not limited to language and direct item translation. Culture, specifically the degree of acculturation and exposure to Western urbanised culture, must be considered as well (Nell, 2000). Having a lower level of educational attainment, living in a rural area, being unfamiliar with the American, Canadian, or British dialect of English, and lack of exposure to aspects of westernised culture and experience may all affect responses to test items (Grieve & Foxcroft, 2013). Two key issues emerge from the literature: (i) the use of objects, words, or pictures that are unfamiliar or have a different name in other cultural contexts, and (ii) the measuring of psychological constructs, emotions, and behaviours that are either absent or understood differently in different cultures. For example, the Boston Naming Test may be culturally biased against English-speaking South Africans. Two studies found that, on average, English-speaking, highly educated South African individuals performed poorly relative to their North American and Australasian counterparts, as they were often unfamiliar with culturally-bound objects (e.g., 'dominoes', 'yoke', 'trellis') presented in pictures (Mendonça, 2010; Thomas et al., 2019). Another example is an 'absurdity' question on the Western Aphasia Battery (i.e.,

"Does it snow in July?"), which may not be suitable for individuals living in the Southern hemisphere (Barratt et al., 2012).

When administering a tool to a population that differs from the original target population, it is important to consider whether the constructs being measured are appropriate for, and are similarly experienced and understood in, that cultural context (Grieve & Foxcroft, 2013). For example, measures of child resilience developed in the global North, such as the Resilience and Youth Development Module, do not consider manifestations of resilience in economically underdeveloped or otherwise marginalised populations, such as issues of food security and physical safety (Ungar & Liebenberg, 2011). Van Eeden and Mantsha's (2007) study on a Venda translation of the 16PF found that the subscale measuring "warmth" had a very low internal consistency, as measured by coefficient alpha ($\alpha < 0.1$). Items describing feelings were generally problematic, suggesting that "warmth" may not manifest in this cultural group as an openness to the expression of emotions. Items measuring expression of feeling might therefore not be suitable for measuring this construct with Venda-speaking individuals. Because the content of test items is tied to the culture in which, and for which, the test was developed, validation studies should also consider, on a qualitative level, whether tools are indeed measuring the constructs they intend to measure in different contexts.

Evaluating the psychometric properties of tools, and translations thereof, is necessary to ascertain both the appropriateness of the language and content of tools for local use. However, many developmental tools commonly used in clinics and for research purposes in South Africa have not been formally validated. Before scores on these measures can be reliably interpreted, psychometric equivalence with the original tools, and appropriate population norms, should be established. Furthermore, translated versions of tools must reliably measure the same constructs as the original tools so that researchers can compare findings across countries and cultures, and so that clinicians can be sure that they are measuring the constructs they intend to measure.

Issues related to tool validity are not limited to those of language and culture. Clinical characteristics of respondents and subjects may also have implications for tool validity. Cognitive and behavioural assessments of children with NDDs, such as Autism Spectrum Disorder (ASD) and Intellectual Disability (ID), may be indicated to inform differential diagnosis and decisions regarding school placements (Paynter & Fothergill, 2015). ASD is characterised by persistent deficits in social communication and interaction, including social reciprocity, non-verbal communication, and understanding social relationships, as well as

repetitive and restricted behaviours, interests, or activities (American Psychiatric Association, 2013). Challenges in social understanding may result in children with ASD being more easily frustrated and more difficult to engage during cognitive assessment (Akshoomoff, 2006). In addition, communication difficulties may impair a child's ability to understand task requirements or respond to instructions. Studies have shown that amongst children with ASD, scores on cognitive tests vary significantly depending on the test administered. Children with ASD obtain significantly lower scores on tests that rely heavily on verbal understanding than on those with lower language demands (Grondhuis et al., 2018; Magiati & Howlin, 2001). A tool that is unbiased should function similarly across different (construct-irrelevant) conditions (Van De Vijver & Leung, 2011). In other words, for a tool to have equivalence, test scores should have the same meaning within and between different groups of people. In the context of children with ASD, it seems that not all cognitive measures are 'equal', and that tests scores may be heavily influenced by the presence of ASD.

Assessment of psychopathology and other comorbid conditions in this population may also present challenges. For example, children and adolescents with NDDs may have reduced capacity for communicating their thoughts and feelings, which may impede the assessment of psychiatric conditions (Grondhuis & Aman, 2012). Studies have shown that verbal ability significantly influences scores on measures of psychopathology (e.g., Brereton et al., 2006). There is also evidence to suggest that psychiatric conditions, such as anxiety, may manifest atypically in children with ASD (Kerns & Kendall, 2012). One study explored the factorial structure of the Multidimensional Anxiety Scale for Children in children with verbal ASD and anxiety disorders, and in a comparison group with age- and gender-matched typically developing children with anxiety disorders. The study did not find equivalent factorial structure in the two groups, suggesting that children and adolescents with ASD may not experience anxiety in the same way as their typically developing counterparts (White et al., 2015). Taken together, these studies suggest that standard measures of psychopathology may not be suitable for use with children who have ASD. These findings also highlight the importance of designing and validating measures of psychopathology for use in this population (Helverschou et al., 2009; Mannion et al., 2014).

Attention-Deficit/Hyperactivity Disorder (ADHD) is an NDD frequently comorbid with other NDDs, such as ASD and ID (Leyfer et al., 2006; Neece et al., 2011). Although these conditions have conceptually different phenotypes, they share overlapping behavioural features and cognitive impairments (Karalunas et al., 2018; Ronald et al., 2014). These overlaps may complicate the assessment of ADHD in children and adolescents with other

NDDs. For example, an ADHD symptom such as "often does not follow through on instructions and fails to finish schoolwork" may indicate inattention, one of the key characteristics of ADHD. However, in the context of ID, the same behaviour may stem from comprehension difficulties or deficits in cognitive skills required to complete a task. Hence, it may be difficult to ascertain whether ADHD symptoms are merely an expression of cognitive impairment or a separate disorder that accounts for the behaviour above and beyond cognitive deficits (Antshel et al., 2006). Consequently, other NDDs may mask or overshadow clinically significant attentional problems or hyperactivity-impulsivity that would lead to a diagnosis of ADHD in typically developing children (Frazier et al., 2001). There is some, limited evidence to suggest that instruments designed to detect ADHD symptoms may not be sensitive to different manifestations of ADHD in children with NDDs. One study evaluating an ADHD rating scale in a sample of verbal children with ASD found that some of the scale items were not adequately tapping into the underlying constructs thought to underpin ADHD, namely inattention and hyperactivity (Yerys et al., 2017). These results suggest that the presence of ASD may influence the ratings of target (ADHD) behaviours.

Overall, then, valid and clinically useful behavioural tools are needed to ensure fair assessment for all, regardless of language background, sociocultural background, or phenotype. Research on the validity of behavioural tools for children in South Africa is still in its infancy. Most published studies have been small, and the psychometric evaluations of the tools cursory. To date, there are only a few systematic reviews summarising the current evidence on behavioural assessment tools in South Africa. Hoosen and colleagues (2018) conducted a scoping review of the Strengths and Difficulties Questionnaire (SDQ) in Africa (most studies came from South Africa), and highlighted that little is known about the cultural and psychometric soundness of the SDQ in this region. With regards to the assessment of comorbid conditions in children with NDDs, most published studies addressing this issue hail from the global North, and typically include children with ASD who do not have communication impairments or comorbid ID (see, e.g., Yerys et al., 2017). There are fewer studies with children who have ASD with associated cognitive impairments and who are non-verbal, as is typical in clinical settings in Africa (Bakare & Munir, 2011; Springer et al., 2013).

Aims

The overarching aim of this dissertation was to investigate the psychometric properties of two behavioural tools; one, a general behavioural screening tool widely used in many different cultural contexts, and the other, an ADHD assessment tool frequently

administered to parents of children with NDDs. This dissertation comprises two distinct studies corresponding to the two behavioural tools.

The first study is a systematic review of the psychometric properties of the Achenbach System of Empirically Based Assessment (ASEBA) forms, the Child Behaviour Checklist (CBCL) and the Youth Self-Report (YSR), in sub-Saharan Africa (SSA). The ASEBA forms are recognised as "gold standard" behavioural assessment tools in both research and clinical contexts worldwide (Achenbach & Rescorla, 2000, 2001; Ivanova, Achenbach, et al., 2007; Ivanova, Dobrean, et al., 2007). I included all sub-Saharan African (SSAn) countries to broaden the scope of the review, thereby enhancing the utility of the findings for researchers and clinicians across the African continent. The ASEBA forms are frequently translated, adapted, and administered in SSAn countries. Despite their widespread use, there are no systematic reviews on the validity, reliability, and cultural appropriateness of the ASEBA forms for the diverse populations and contexts of SSA. This review sought to address this gap by evaluating the use and reported psychometric properties of the ASEBA forms in this region. The specific aims of the systematic review were:

1. to identify and collate all studies that used the ASEBA forms with SSAn participants,
2. to describe the use of the ASEBA forms across SSA, including the use of translations and adaptations thereof,
3. to systematically evaluate the reported psychometric properties of the different forms and subscales, and finally,
4. to make recommendations regarding the use of the ASEBA forms in SSA based on available evidence.

The second study evaluates the validity of the Swanson, Nolan, and Pelham ADHD Rating Scale (4th edition; SNAP-IV) in a sample of South African children with NDDs. The SNAP-IV, developed in the United States, is a parent-report rating scale designed to measure behaviours associated with ADHD (Swanson et al., 2012). I selected an ADHD rating scale as ADHD is a common comorbidity in children with other NDDs, including ASD and ID (Leyfer et al., 2006; Neece et al., 2011; Simonoff et al., 2007). The SNAP-IV is administered in clinical and school settings in South Africa, but very few local studies have published on it and reported its psychometric properties. To date, only a few studies internationally have investigated the validity of the SNAP-IV in children with ID, and, to our knowledge, none with children who have ASD. The specific aims of the present study were:

1. to investigate the structural validity of the SNAP-IV using confirmatory factor analysis (CFA) techniques,

2. to estimate the internal consistency reliability of the SNAP-IV subscales, ‘Inattention’ and ‘Hyperactivity-Impulsivity’, and
3. to evaluate the construct validity, specifically, the concurrent, convergent, and discriminant validity of the SNAP-IV subscales by correlating subscale scores with those of another similar tool.

Outline of the Chapters to Follow

Chapter 2

Chapter 2 presents a manuscript of a systematic review of studies from SSA that use the ASEBA forms. The review identifies and describes all studies that report psychometric properties of the ASEBA forms with SSAn samples. The review also conducts an in-depth evaluation of the reported measurement properties, identifies gaps in the literature, and makes recommendations on the use of the ASEBA forms in SSAn contexts.

Chapter 3

Chapter 3 describes the methods and context for the study presented in Chapter 4. Although Chapter 4 includes its own methods section, the word restrictions of the journal did not allow for adequate contextual detail. This chapter also details the statistical techniques employed for the psychometric analyses.

Chapter 4

Chapter 4 presents a manuscript of a study evaluating the validity and internal consistency of the SNAP-IV in a sample of South African children with NDDs. I discuss the implications of the findings for clinical contexts in South Africa.

Chapter 5

In Chapter 5, I discuss and synthesise the findings of the two studies reported in Chapters 2 and 4. I elaborate further on limitations of each study, as well as the specific implications and significance of the findings. I conclude that these data highlight the importance of using valid and reliable measures in different contexts. These data underscore the need for comprehensive psychometric evaluation of behavioural assessment tools. I conclude by suggesting directions for future research.

Author Contributions to Included Manuscripts

This dissertation includes two manuscripts that have been submitted to peer-reviewed journals. I describe my contributions to each manuscript below.

1. *Psychometric properties of the ASEBA Child Behaviour Checklist and Youth Self-Report in sub-Saharan Africa: A systematic review*, **Michal R. Zieff**, Claire Fourie, Michelle Hoogenhout, & Kirsten A. Donald. *Under review: Acta Neuropsychiatrica.*

I developed a search methodology with input from my co-authors, searched the selected databases, and saved all citation records to a reference manager (EndNote X9). CF and I independently conducted a three-stage systematic review of all records to check for relevance and eligibility. We then compared findings and discussed any discrepancies until we reached a consensus. I then extracted, summarised, and analysed the data from the included studies. I wrote a full draft of the manuscript. All authors reviewed the manuscript and made intellectual contributions in their respective areas of expertise. Finally, I integrated all authors' feedback and submitted the finalised manuscript to a peer-reviewed journal for consideration.

2. *Validity of the SNAP-IV for ADHD assessment in South African children with neurodevelopmental disorders*, **Michal R. Zieff**, Michelle Hoogenhout, Emma Eastman, Björn U. Christ, Alice Galvin, Victoria de Menil, Amina Abubakar, Charles R. Newton, Elise Robinson, & Kirsten A. Donald. *Under review: Journal of Autism and Developmental Disorders.*

I conceptualised the study, collected, analysed and interpreted the data, and wrote the first full draft of the manuscript. KAD, ER, AA, CRN, and VdM made significant contributions to the conceptualisation and design of the umbrella study (The NeuroDEV Study; de Menil et al., 2019). AG and EE were involved in the conceptualisation and logistical operations of the study. They were also responsible for the management and quality control of the data. BUC and EE were instrumental in the acquisition of data. They were also involved in selecting and preparing behavioural measures for the study. MH provided expert-level statistical support and checked that the analyses and interpretation of the data were sound. Finally, KAD and MH provided in-depth feedback on the interpretation of the results and conclusions, as well as the significance of the results in clinical and research contexts. All authors reviewed the manuscript and gave suggestions for revision. I integrated all authors' contributions and submitted the finalised manuscript to a peer-reviewed journal for consideration.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. University of Vermont.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. University of Vermont.

Akshoomoff, N. (2006). Use of the Mullen Scales of Early Learning for the assessment of young children with Autism Spectrum Disorders. *Child Neuropsychology*, *12*(4–5), 269–277. https://doi.org/10.1080/09297040500473714

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

Antshel, K. M., Phillips, M. H., Gordon, M., Barkley, R., & Faraone, S. V. (2006). Is ADHD a valid disorder in children with intellectual delays? *Clinical Psychology Review*, *26*(5), 555–572. https://doi.org/10.1016/j.cpr.2006.03.002

Bakare, M. O., & Munir, K. M. (2011). Excess of non-verbal cases of autism spectrum disorders presenting to orthodox clinical practice in Africa – a trend possibly resulting from late diagnosis and intervention. *South African Journal of Psychiatry*, *17*(4), 118–120. https://doi.org/10.4102/sajpsychiatry.v17i4.295

Barratt, J., Khoza-Shangase, K., & Msimang, K. (2012). Speech-language assessment in a linguistically diverse setting: Preliminary exploration of the possible impact of informal 'solutions' within the South African context. *South African Journal of Communication Disorders*, *59*(1), 34–44. https://doi.org/10.4102/SAJCD.V59I1.20

Brereton, A. V., Tonge, B. J., & Einfeld, S. L. (2006). Psychopathology in children and adolescents with Autism compared to young people with Intellectual Disability. *Journal of Autism and Developmental Disorders*, *36*, 863–870. https://doi.org/10.1007/s10803-006-0125-y

De Kock, F., Kanjee, A., & Foxcroft, C. (2013). Cross-cultural test adaptation, translation and tests in multiple languages. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed., pp. 102–125). Oxford University Press.

de Menil, V., Hoogenhout, M., Kipkemoi, P., Kamuya, D., Eastman, E., Galvin, A., Mwangasha, K., de Vries, J., Kariuki, S. M., Murugasen, S., Mwangi, P., Singh, I., Stein, D. J., Abubakar, A., Newton, C. R., Donald, K. A., & Robinson, E. (2019). The NeuroDev Study: Phenotypic and genetic characterization of neurodevelopmental disorders in Kenya and South Africa. *Neuron*, *101*(1), 15–19. https://doi.org/10.1016/j.neuron.2018.12.016

Fernald, L. C. H., Kariger, P., Engle, P. L., & Raikes, A. (2009). *Examining early child development in low-income countries: A toolkit for the assessment of children in the first five years of life*. The World Bank. https://elibrary.worldbank.org/doi/abs/10.1596/28107

Frazier, J. A., Biederman, J., Bellordre, C. A., Garfield, S. B., Geller, D. A., Coffey, B. J., & Faraone, S. V. (2001). Should the diagnosis of Attention-Deficit/ Hyperactivity disorder be considered in children with Pervasive Developmental Disorder? *Journal of Attention Disorders*, *4*(4), 203–211. https://doi.org/10.1177/108705470100400402

Frick, P. J., Barry, C. T., & Kamphaus, R. W. (2020). *Clinical assessment of child and adolescent personality and behavior* (4th ed.). Springer.

Grieve, K. W., & Foxcroft, C. (2013). Factors affecting assessment results. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed., pp. 282–297). Oxford University Press.

Grieve, K. W., & Van Eeden, R. (2010). A preliminary investigation of the suitability of the WAIS-III for Afrikaans-speaking South Africans. *South African Journal of Psychology*, *40*(3), 262–271. https://doi.org/10.10520/EJC98598

Grondhuis, S. N., & Aman, M. G. (2012). Assessment of anxiety in children and adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *6*(4), 1345–1365. https://doi.org/10.1016/J.RASD.2012.04.006

Grondhuis, S. N., Lecavalier, L., Arnold, L. E., Handen, B. L., Scahill, L., McDougle, C. J., & Aman, M. G. (2018). Differences in verbal and nonverbal IQ test scores in children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *49*, 47–55. https://doi.org/10.1016/J.RASD.2018.02.001

Helverschou, S. B., Bakken, T. L., & Martinsen, H. (2009). The Psychopathology in Autism Checklist (PAC): A pilot study. *Research in Autism Spectrum Disorders*, *3*(1), 179–195. https://doi.org/10.1016/j.rasd.2008.05.004

Hoosen, N., Davids, E. L., de Vries, P. J., & Shung-King, M. (2018). The Strengths and Difficulties Questionnaire (SDQ) in Africa: A scoping review of its application and validation. *Child and Adolescent Psychiatry and Mental Health*, *12*(1), Article 6. https://doi.org/10.1186/s13034-017-0212-1

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, *3*, 29–51. https://doi.org/10.1146/annurev.clinpsy.3.022806.091419

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Dumenci, L., Almqvist, F., Bilenberg, N., Bird, H., Broberg, A. G., Dobrean, A., Döpfner, M., Erol, N., Forns, M., Hannesdottir, H., Kanbayashi, Y., Lambert, M. C., Leung, P., Minaei, A., Mulatu, M. S., Novik, T., … Verhulst, F. C. (2007). The generalizability of the Youth Self-Report syndrome structure in 23 societies. *Journal of Consulting and Clinical Psychology*, *75*(5), 729–738. https://doi.org/10.1037/0022-006X.75.5.729

Ivanova, M. Y., Dobrean, A., Dopfner, M., Erol, N., Fombonne, E., Fonseca, A. C., Frigerio, A., Grietens, H., Hannesdottir, H., Kanbayashi, Y., Lambert, M., Achenbach, T. M., Larsson, B., Leung, P., Liu, X., Minaei, A., Mulatu, M. S., Novik, T. S., Oh, K. J., … Chen, W. J. (2007). Testing the 8-syndrome structure of the Child Behavior Checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology*, *36*(3), 405–417. https://doi.org/10.1080/15374410701444363

Karalunas, S. L., Hawkey, E., Gustafsson, H., Miller, M., Langhorst, M., Cordova, M., Fair, D., & Nigg, J. T. (2018). Overlapping and distinct cognitive impairments in Attention-Deficit/Hyperactivity and Autism Spectrum Disorder without Intellectual Disability. *Journal of Abnormal Child Psychology*, *46*(8), 1705–1716. https://doi.org/10.1007/s10802-017-0394-2

Kerns, C. M., & Kendall, P. C. (2012). The presentation and classification of anxiety in Autism Spectrum Disorder. *Clinical Psychology: Science and Practice*, *19*(4), 323–347. https://doi.org/10.1111/CPSP.12009

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., Tager-Flusberg, H., & Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, *36*(7), 849–861. https://doi.org/10.1007/s10803-006-0123-0

Magiati, I., & Howlin, P. (2001). Monitoring the progress of preschool children with autism enrolled in early intervention programmes. *Autism*, *5*(4), 399–406. https://doi.org/10.1177/1362361301005004005

Mannion, A., Brahm, M., & Leader, G. (2014). Comorbid psychopathology in Autism Spectrum Disorder. *Review Journal of Autism and Developmental Disorders*, *1*, 124–134. https://doi.org/10.1007/s40489-014-0012-y

Mendonça, J. (2010). *Performance of English, Zulu and Sotho students on the Boston Naming test: An investigation into the items responsible for cultural bias* [Unpublished masters dissertation]. University of the Witwatersrand. http://hdl.handle.net/10539/11440

Neece, C. L., Baker, B. L., Blacher, J., & Crnic, K. A. (2011). Attention-deficit/hyperactivity disorder among children with and without intellectual disability: An examination across time. *Journal of Intellectual Disability Research*, *55*(7), 623–635. https://doi.org/10.1111/j.1365-2788.2011.01416.x

Nell, V. (2000). *Cross-cultural neuropsychological assessment: Theory and practice*. Erlbaum.

Nezafat Maldonado, B., Chandna, J., & Gladstone, M. (2019). A systematic review of tools used to screen and assess for externalising behaviour symptoms in low and middle income settings. *Global Mental Health*, *6*, Article e13. https://doi.org/10.1017/gmh.2019.11

Paynter, J., & Fothergill, H. (2015). Conducting developmental and cognitive assessments with young children with Autism Spectrum Disorders. *Australian Clinical Psychologist*, *1*(2), 14–18.

Ronald, A., Larsson, H., Anckarsäter, H., & Lichtenstein, P. (2014). Symptoms of autism and ADHD: A swedish twin study examining their overlap. *Journal of Abnormal Psychology*, *123*(2), 440–451. https://doi.org/10.1037/A0036088

Sharp, C., Venta, A., Marais, L., Skinner, D., Lenka, M., & Serekoane, J. (2014). First evaluation of a population-based screen to detect emotional-behavior disorders in orphaned children in sub-Saharan Africa. *AIDS and Behavior*, *18*(6), 1174–1185. https://doi.org/10.1007/s10461-014-0739-6

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G. L., 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. https://doi.org/10.1080/13803390490510824

Simonoff, E., Pickles, A., Wood, N., Gringras, P., & Chadwick, O. (2007). ADHD symptoms in children with mild intellectual disability. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(5), 591–600. https://doi.org/10.1097/chi.0b013e3180323330

Springer, P. E., Van Toorn, R., Laughton, B., & Kidd, M. (2013). Characteristics of children with pervasive developmental disorders attending a developmental clinic in the Western Cape Province, South Africa. *South African Journal of Child Health*, *7*(3), 95–99. https://doi.org/10.7196/sajch.530

Statistics South Africa. (2012). *Census 2011: Census in brief*. http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., Clevenger, W., Wasdell, M., McCleary, R., Lakes, K., & Wigal, T. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN Rating Scales. *The International Journal of Educational and Psychological Assessment*, *10*(1), 51–70.

Sweetland, A. C., Belkin, G. S., & Verdeli, H. (2014). Measuring depression and anxiety in sub-Saharan Africa. *Depression and Anxiety*, *31*(3), 223–232. https://doi.org/10.1002/da.22142

Thomas, K. G. F., Baerecke, L., Pan, C. Y., & Ferrett, H. L. (2019). The Boston Naming Test-South African Short Form, Part I: Psychometric properties in a group of healthy English-speaking university students. *African Journal of Psychological Assessment*, *1*(0), Article a15. https://doi.org/10.4102/AJOPA.V1I0.15

Ungar, M., & Liebenberg, L. (2011). Assessing resilience across cultures using mixed methods: Construction of the Child and Youth Resilience Measure. *Journal of Mixed Methods Research*, *5*(2), 126–149. https://doi.org/10.1177/1558689811400607

Van De Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17–45). Cambridge University Press.

Van Eeden, R., & Mantsha, T. R. (2007). Theoretical and methodological considerations in the translation of the 16PF5 into an African language. *South African Journal of Psychology*, *37*(1), 62–81. https://doi.org/10.1177/008124630703700105

Van Widenfelt, B. M., Treffers, P. D. A., De Beurs, E., Siebelink, B., Koudijs, E., & Siebelink, B. M. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child & Family Psychology Review*, *8*(2), 135–147. https://doi.org/10.1007/s10567-005-4752-1

White, S. W., Lerner, M. D., McLeod, B. D., Wood, J. J., Ginsburg, G. S., Kerns, C., Ollendick, T., Kendall, P. C., Piacentini, J., Walkup, J., & Compton, S. (2015). Anxiety in youth with and without Autism Spectrum Disorder: Examination of factorial equivalence. *Behavior Therapy*, *46*(1), 40–53. https://doi.org/10.1016/J.BETH.2014.05.005

Yerys, B. E., Nissley-Tsiopinis, J., de Marchena, A., Watkins, M. W., Antezana, L., Power, T. J., & Schultz, R. T. (2017). Evaluation of the ADHD Rating Scale in youth with Autism. *Journal of Autism and Developmental Disorders*, *47*(1), 90–100. https://doi.org/10.1007/s10803-016-2933-z

Zieff, M. R. (2017). *A psychometric evaluation of a Xhosa translation of the SA-WASI Vocabulary subtest* [Unpublished honours project]. University of Cape Town. http://www.uct.ac.za/sites/default/files/image_tool/images/117/Logos/thesis/MichalZieff_FinalHonsResProject_2017.pdf

In Chapter 2, I present a systematic review of the psychometric properties of the ASEBA forms in sub-Saharan Africa. I developed the search methodology for the review and conducted the database searches. I reviewed all records, and then extracted, summarised, and analysed data from the eligible studies. I wrote the full manuscript with input from all my co-authors and submitted the finalised draft to a peer-reviewed journal. The chapter is presented in the format of the submitted manuscript.

CHAPTER 2

Psychometric Properties of the ASEBA Child Behaviour Checklist and Youth Self-Report in Sub-Saharan Africa: A Systematic Review

Michal R. Zieff[1], Claire Fourie[1], Michelle Hoogenhout[1,2], & Kirsten A. Donald[1,3]

[1]Division of Developmental Paediatrics, Department of Paediatrics and Child Health, University of Cape Town, Red Cross War Memorial Children's Hospital, Klipfontein Road, Rondebosch 7700, Cape Town, South Africa
[2]Galvanize, 44 Tehama Street, San Francisco, CA 94501, USA
[3]Neuroscience Institute, University of Cape Town, Cape Town, South Africa

Acknowledgements

We thank Mrs Mary Shelton from the University of Cape Town's Health Sciences Library for her advice and assistance with the systematic review search strategy.

Financial Support

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Statement of Interest

None.

Abstract

Objective: Behavioural screening tools may be used to identify at-risk children in resource-limited settings in sub-Saharan Africa. The ASEBA forms (Child Behaviour Checklist and Youth Self-Report) are frequently translated and adapted for use in sub-Saharan African populations, but little is known about their measurement properties in these contexts. *Methods:* We conducted a systematic review of all published journal articles that used the ASEBA forms with sub-Saharan African samples. We evaluated the reported psychometric properties, as well as the methodological quality of the psychometric evaluations, using COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) guidelines. *Results:* Fifty-eight studies reported measurement properties of the ASEBA forms. Most studies came from Southern (n = 29, 50%) or East African (n = 25, 43%) countries. Forty-nine studies (84%) used translated versions of the tool, but details regarding the translation process, if available, were often sparse. Most studies (n = 47, 81%) only reported internal consistency (using coefficient alpha) for one or more subscale. The methodological quality of the psychometric evaluations ranged from 'very good' to 'inadequate' across all measurement properties, except for internal consistency. *Conclusion:* There is limited good quality psychometric evidence available for the ASEBA forms in sub-Saharan Africa. We recommend (i) implementing a standardised procedure for conducting and reporting translation processes, and (ii) conducting more comprehensive psychometric evaluations of the translated versions of the tools.

Keywords: ASEBA, Child Behaviour Checklist, Youth Self Report, psychometric, sub-Saharan Africa

Summations

1. The search identified 58 studies that administered the ASEBA Child Behaviour Checklist or Youth Self-Report forms to sub-Saharan African participants and reported at least one psychometric property.
2. Most studies reported only coefficient alpha as a measure of internal consistency. Only nine studies, all from East African countries, were specifically focused on the psychometric properties of the ASEBA forms.
3. There is some evidence to support the structural validity and internal consistency of the ASEBA forms. Evidence concerning the content validity, criterion validity, and reliability of the ASEBA forms is limited.

Considerations

1. Although most studies used translated versions of the ASEBA forms, very few reported details on the translation and adaptation process. This limited our ability to evaluate the validity of the translation process.
2. Inconsistencies and large variation (not just in measurement properties, but in country, language as well as other factors) meant that we could not quantitatively pool the results and arrive at reasonable conclusions.
3. Limitations and inconsistencies meant that we were not able to make definitive recommendations regarding the use of the ASEBA forms in sub-Saharan Africa based on the available evidence. More comprehensive, good quality psychometric studies are needed.

Behavioural and emotional problems represent a significant health burden amongst children and adolescents in sub-Saharan Africa (SSA; M. A. Cortina et al., 2012; Kusi-Mensah et al., 2019; Mulatu, 1995; Ndetei et al., 2016). As specialist child mental health services in SSA are limited (Jenkins et al., 2010), screening tools, such as the Achenbach System of Empirically Based Assessment (ASEBA) forms (Achenbach & Rescorla, 2000, 2001), are often administered in community settings to detect common childhood emotional and behavioural problems (Hoosen et al., 2018). Screening tools typically do not require specialist training and are generally quick and easy to administer and score, making them particularly advantageous in low- and middle-income settings (Sharp et al., 2014). Despite their widespread use, there are no systematic reviews on the validity, reliability, and cultural appropriateness of the ASEBA forms in the diverse populations and contexts of SSA. This review seeks to address that gap by evaluating the use and reported psychometric properties of the ASEBA forms in SSA.

The majority of behavioural screening tools used in SSA are developed in North America or Europe and are generally well established in these regions (Fernald et al., 2009; Sharp et al., 2014; Sweetland et al., 2014). Using such tools in SSA is often more efficient and feasible compared to developing new tools locally, and also enables comparison of findings cross-culturally (Van Widenfelt et al., 2005). However, tools developed in the global north are generally designed for direct application to English-speaking individuals from Western, urbanised, populations (Nezafat Maldonado et al., 2019). It follows that using a tool with individuals who are not English-speaking, or who are from cultures that differ substantially from that of the original target population, may present issues for both administration and interpretation (Sweetland et al., 2014).

There are several challenges associated with translating a tool from English into another language (De Kock et al., 2013; Fernald et al., 2009). For example, an English word or phrase may not have a linguistic equivalent in the target language, or, if an equivalent word or phrase does exist, it may not form part of the vernacular of the target population. A direct translation may also have a slightly different or ambiguous meaning in the target language (Van Widenfelt et al., 2005). Many African languages do not have established terms to describe specific mental illnesses, emotions, or personality traits (Atilola, 2015; Van Eeden & Mantsha, 2007). Poor translations of items that measure psychological constructs may therefore introduce bias, which may, in turn, compromise the validity and reliability of the scores derived from the tool.

It is also important to consider whether constructs being measured by a tool are relevant and understood in the same way in different cultures (i.e., construct equivalence; Van De Vijver & Leung, 2011). This is applicable even when a tool is administered in its original English. A South African study conducted a pilot test to evaluate the cultural appropriateness of items on the Child Behaviour Checklist (CBCL) in English and in two other South African languages (LeCroix et al., 2020; Palin et al., 2009). Feedback from participants led to the removal of the item "sets fires", intended to measure rule-breaking behaviour. In this context, it is likely that setting fires (e.g., for cooking) is commonplace amongst children and adolescents as part of daily life, and so participants interpreted the item in this way, instead of the intended interpretation (i.e., setting a fire with intent to cause harm or damage). Hence, establishing linguistic and construct equivalence prior to using a tool outside of its original context is critical to ensure that the tool is measuring what it intends to measure.

Measurement, or psychometric, properties (e.g., validity and reliability) are not properties of the tool itself, but are characteristics of the data derived from the tool in a specific context (Zumbo & Chan, 2014). Most applied research studies conduct rudimentary psychometric evaluations of scores obtained from psychological tools (Dima, 2018; Flake et al., 2017; Vacha-Haase & Thompson, 2011). The result is the use of tools, including behavioural screening tools, without sufficient evidence to support the validity and reliability of the scores derived from the tools in a given context. Scores generated from such a tool may not accurately reflect the 'true scores' of the respondents (De Kock et al., 2013). This may, in turn, increase the risk of misdiagnosis, which has implications for referral and the provision of appropriate interventions. Hence, until psychometric equivalence of a behavioural screening tool is established, results should be interpreted with caution.

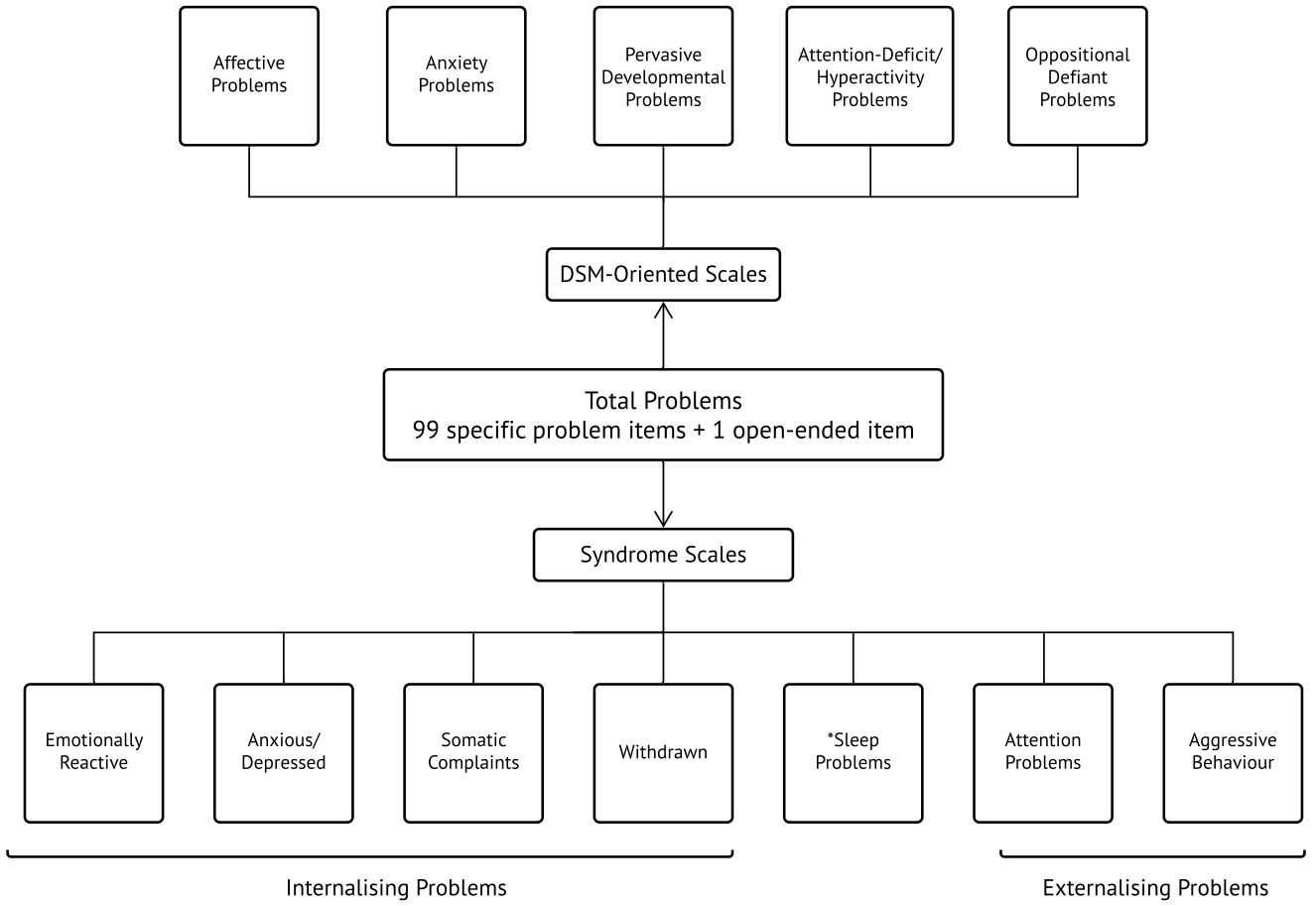
The ASEBA forms, developed in the United States, are currently used as screening tools for clinical and research purposes in SSA. The ASEBA forms are designed to quickly and effectively measure maladaptive behaviours in children and adolescents. One major advantage of the ASEBA forms is that data can be obtained from multiple informants (i.e., caregiver, teacher, and self-report), allowing for a comprehensive overview of the child's behaviour in different contexts. The most recent version of the Preschool Forms include the parent-report Child Behaviour Checklist for ages 1.5-5 (CBCL/1.5-5) and the Caregiver-Teacher Report Form for Ages 1.5-5 (C-TRF; Achenbach & Rescorla, 2000). The School-Age Forms include the parent-report Child Behaviour Checklist for Ages 6-18 (CBCL/6-18), the Teacher's Report Form (TRF), and the Youth Self Report (YSR; Achenbach & Rescorla,

2001). In this review, we refer to the parent report forms collectively as the 'CBCL', but when referring to a specific age form, we use the corresponding ASEBA abbreviation (e.g., CBCL/1.5-5 or CBCL/6-18).

The forms are presented as lists of items describing a range of behaviours, (e.g., "avoids looking others in the eye", "has trouble getting to sleep"). Respondents indicate their agreement with the items by selecting either "not true" (scored as a 0), "somewhat or sometimes true" (scored as a 1) or "very true/often true" (scored as a 2). These scores are summed to provide a Total Problems score, where higher scores indicate the presence of more problem behaviours. Items are grouped into syndrome scales, which are further grouped into two broad band scales (Internalising Problems and Externalising Problems, see Figures 1 and 2). Items are also grouped into DSM-oriented scales (see Figures 1 and 2), aligned with diagnostic criteria for a number of disorders specified in the fifth edition of the Diagnostic and Statistical Manual for Mental Disorders (DSM-5; American Psychiatric Association, 2013).

Figure 1

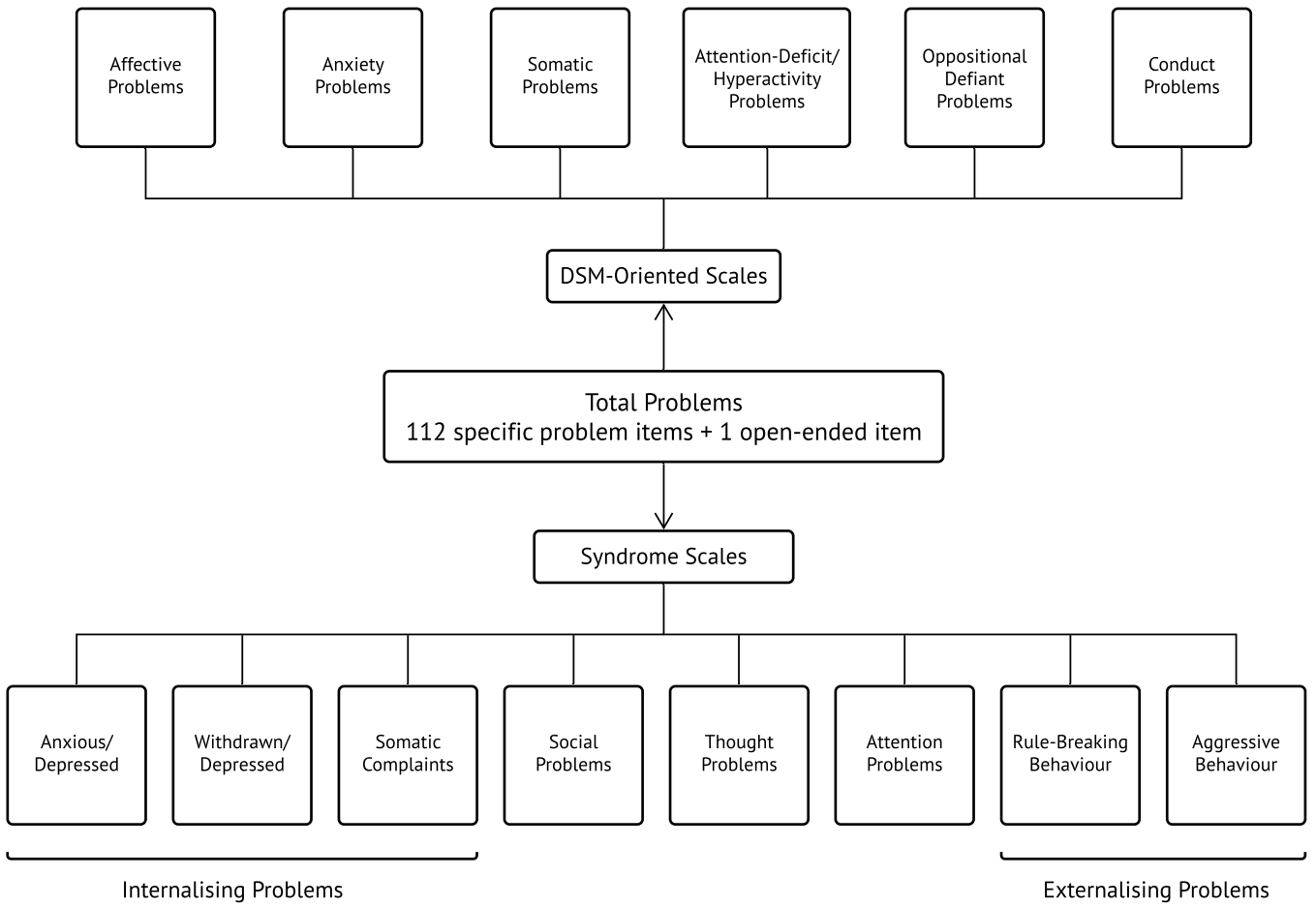
Structure of the ASEBA Preschool Forms (CBCL/1.5-5 and C-TRF): Syndrome Scales and DSM-Oriented Scales



Note. ASEBA = Achenbach System of Empirically Based Assessment, CBCL/1.5-5 = Child Behaviour Checklist for Ages 1.5-5, C-TRF = Caregiver-Teacher Report Form, DSM = Diagnostic and Statistical Manual of Mental Disorders. *Sleep Problems syndrome scale is featured only in the CBCL/1.5-5, not the C-TRF. Figure created with Lucidchart.com.

Figure 2

Structure of the ASEBA School-Age Forms (CBCL/6-18, YSR, and TRF): Syndrome Scales and DSM-Oriented Scales



Note. ASEBA = Achenbach System of Empirically Based Assessment, CBCL/6-18 = Child Behaviour Checklist for Ages 6-18, YSR = Youth Self Report, TRF = Teacher's Report Form, DSM = Diagnostic and Statistical Manual of Mental Disorders. Figure created with Lucidchart.com.

It is not clear where or how the ASEBA forms are used in SSA, or to what extent the scores from the ASEBA forms have been evaluated for their validity, reliability, and cultural appropriateness for the diverse populations and contexts in this region. This study had four primary objectives, namely (i) to collate all studies that used the ASEBA forms with sub-Saharan African (SSAn) participants, (ii) to describe the use of the ASEBA forms across SSA (including the use of translations), (iii) to evaluate the reported psychometric properties of the scores of different forms and subscales, and finally, (iv) to make recommendations regarding the use of the ASEBA forms in SSA based on available evidence.

Methods

We searched PubMed, EBSCO (APA PsycInfo, APA PsycArticles, ERIC, Academic Search Premier, Health Source: Nursing/Academic Edition, Africa-Wide Information, CINAHL), Scopus, and Google Scholar databases. In addition, we searched ProQuest and the University of Cape Town's (UCT) library database for relevant dissertations/theses, book chapters, and conference abstracts. A detailed overview of the search strategy is presented in Table S1 in the digital supplement.

Search Strategy

For all database searches, the ASEBA terms "child behaviour checklist" OR "child behavior checklist" OR CBCL OR "Youth Self-Report" OR "Teacher's Report Form" were added as the first line of the search. Although most publications used the American spelling ("behavior"), the British spelling was included as an option so as not to exclude articles from journals that use British spelling. A preliminary search including only the ASEBA terms revealed a number of journal articles that referred to the CBCL as the "Children's Behaviour Checklist", "Child Behaviour Check List", or "CBC" to refer to the same tool. However, there was no substantial difference in the number of search results when these variants were included in the final search terms. We conducted another trial search without inverted commas around "Child Behaviour Checklist" to check if authors were using the tool name more loosely. However, many more results showed up and most of them were not meaningful (i.e., included other tools with the word 'checklist' in their names). We did not include the abbreviations for the Youth Self-Report (YSR) or the Teacher's Report Form (TRF). Including "YSR" did not substantially affect the number of results and including "TRF" generated too many irrelevant results.

The SSAn search terms were adapted from Pienaar and colleagues (2011) and the list of SSAn countries from the United Nation's Standard Country or Area Codes for Statistical

Use (United Nations Statistics Division, 1998). We excluded surrounding islands and territories (e.g., Madagascar, Comoros) from the SSAn search terms, as we were primarily interested in continental SSAn countries. For four countries, we included alternative names (e.g., "Ivory Coast" in addition to "Côte d'Ivoire"). After conducting preliminary searches with the ASEBA and SSA terms described above, we noticed that many results were coming up with African-American samples. Hence, for all searches, we narrowed the search by excluding the following terms: "African-American" OR "African American".

We did not include any psychometrics-related words or terms (e.g., validity, Cronbach's alpha) because of (i) the broad range of psychometric-related terms, and (ii) inconsistencies in indexing and reporting of psychometric properties. As it was more likely that a study utilised the CBCL as an outcome measure rather than the tool itself being the subject of the study, no Medical Subject Headings (MeSH terms) were included in the search terms. For the same reason, "all/full text" fields were selected for all lines of the searches, except for one database where it yielded too many results (see Table S1 in the digital supplement). No coverage dates (i.e., year limits) were selected in any database search, nor was the type of publication. All records were saved to a reference library (Endnote X9), after which duplicate records were removed.

Inclusion Criteria

A study was eligible to be included in the final analysis, if:

i. The study was written in English.
ii. The study reported original findings.
iii. The study sample (or at least a portion of the sample) was from a SSAn country. Immigrants and refugees currently living outside SSA were eligible if the study reported specific data for the SSAn participants. For immigrants/refugees, either the child or at least one parent/caregiver had to have been born in a SSAn country.
iv. The study used an ASEBA form (any form or any version) in its standard format (i.e., as per the official instructions in the ASEBA manual), and reported the data derived from the tool. Minor adaptations to the tool (e.g., excluding items due to cultural inappropriateness etc.) were acceptable, as long as the modifications were clearly specified and justified, and the tool was still recognisable as an ASEBA form.
v. The study reported psychometric properties (e.g., validity, reliability) of the ASEBA form for the study sample. Inherent in this criterion is the exclusion of case studies.

Screening and Review Process

Two of the authors (MRZ and CF) independently screened and reviewed all records for eligibility. As the ASEBA forms were generally not the subject of the study (and therefore did not appear in the title or abstract), the full text of each article was scanned or read at each stage of the review process until relevance or eligibility (or lack thereof) became clear. At each stage, the reviewers identified and discussed any discrepancies until a consensus was reached. In the event of an impasse, the reviewers presented the article in question to the fourth author (KAD), who made the final decision.

The review comprised three distinct stages:

i. A brief screening of all full-text studies to check for relevance. Did the study include a SSAn sample and use an ASEBA form?

ii. A more thorough screening of the relevant studies to check for eligibility. Did the study describe the SSAn sample (or sub-sample) and were there specific data for those participants? Was the ASEBA form used in the standard way? We excluded studies if the description of the sample or the country of origin was vague or if there were no specific data pertaining to the SSAn participants. We also excluded studies that used an ASEBA form in a non-standard way, as we wanted to reasonably compare the psychometric properties across studies. At this stage, the reviewers scanned the reference lists of relevant articles to look for other literature that did not appear in the original search results.

iii. A review of the studies identified at the second stage to determine if the study reported psychometric properties of the tool. Any psychometric analyses were acceptable, so long as the statistics were for the study sample (i.e., not from another study or the tool manual). We included studies that met these criteria in the final analysis.

We extracted and summarised key information from the included studies, such as details related to the sample, the country of origin, the ASEBA versions administered, the language(s) of administration, any translation or adaptation processes, and the psychometric analyses conducted. If any details were missing from an article or were unclear, we contacted the corresponding author for clarification or, if applicable, referred to other articles related to the same umbrella study. In the event that the corresponding author did not respond after two attempts to contact them, we noted the uncertainty in our records.

We then evaluated the psychometric properties of the ASEBA forms with reference to COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments)

criteria for good measurement properties (Prinsen et al., 2018). COSMIN describes three phases of psychometric evaluation. The first phase involves investigating a tool's content validity, that is, the extent to which the content of the tool adequately reflects the construct being measured. Specifically, COSMIN recommends that tools be relevant, comprehensive, and comprehensible with respect to the construct of interest and the target population (Terwee, Prinsen, Chiarotto, Westerman, et al., 2018). The second phase concerns evaluating the internal structure of the tool, including structural validity, internal consistency, and cross-cultural validity (measurement invariance). The third phase involves evaluating the remaining measurement properties, including reliability, measurement error, criterion validity, and hypothesis testing for construct validity, including concurrent, convergent, divergent, and known-groups validity (Mokkink, Prinsen, et al., 2018). Each measurement property is rated on a three-point scale of 'sufficient', 'indeterminate', or 'insufficient'.

We also assessed the methodological quality of the studies using the COSMIN Risk of Bias checklist (Mokkink, de Vet, et al., 2018). COSMIN utilises a four-point rating system to grade each measurement property in a study as 'very good', 'adequate', 'doubtful' or 'inadequate'. The overall rating of the quality of each measurement property is determined by taking the lowest rating of any standards corresponding to that property.

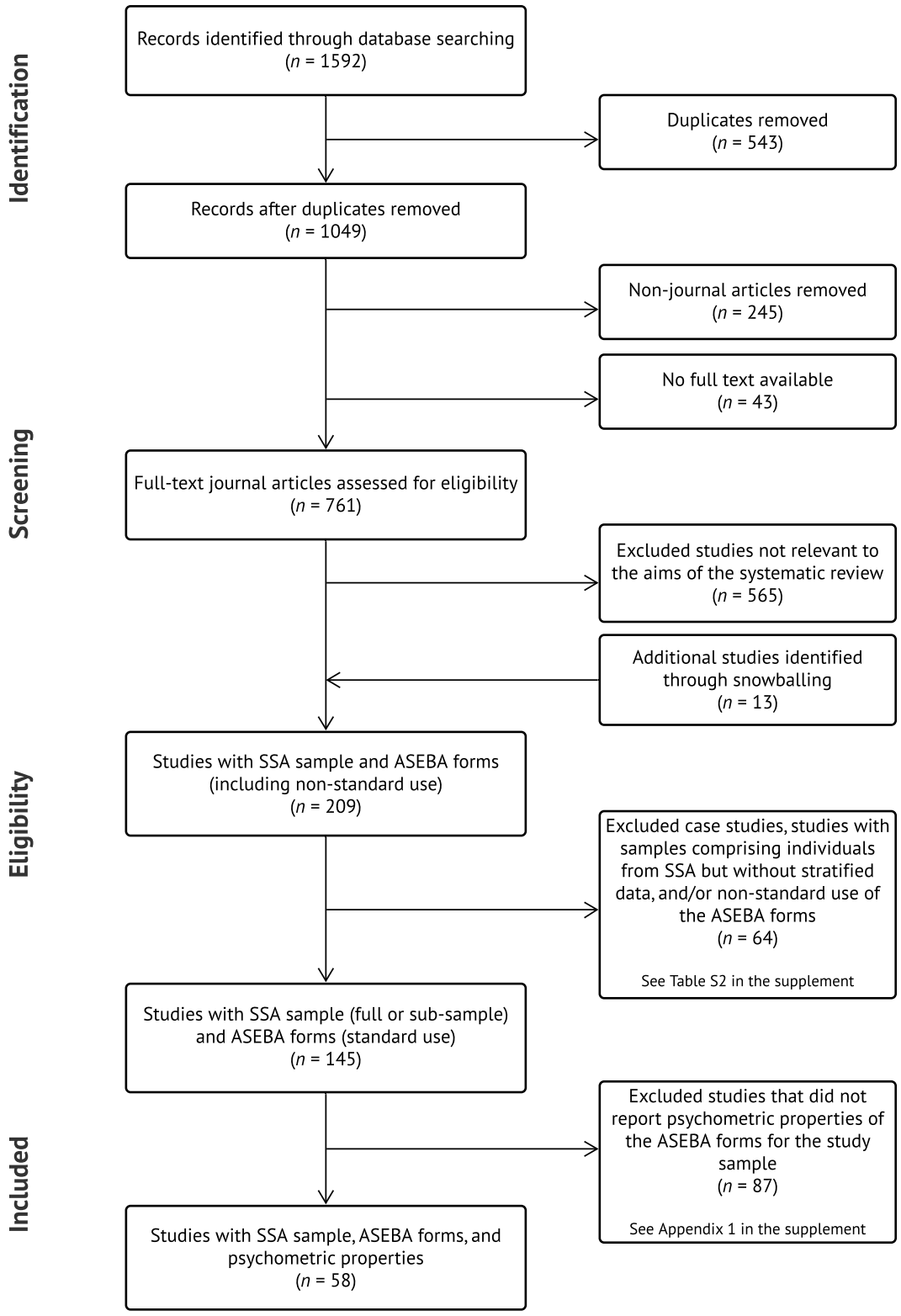
Results

Search Results

A flow diagram adapted from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement (Moher et al., 2010), details the number of studies included and excluded at each stage of the review process (Figure 3).

Figure 3

PRISMA Diagram Adapted from Moher et al. (2010) Outlining the Number of Records Included and Excluded at Each Phase of the Systematic Review



Note. PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses, ASEBA = Achenbach System of Empirically Based Assessment, SSA = sub-Saharan Africa. Figure created with Lucidchart.com.

During the search for manuscripts, it became apparent that full-text versions of many "grey" references, including conference abstracts, poster presentations, dissertations/theses, and other unpublished work, were difficult to access. We decided to randomly sample 20 records from each category of grey literature to check for relevance. The aim of this test was to determine if it was worth pursuing the full-text search for these records. Out of the 20 conference abstracts/presentations sampled, only one was relevant to the aims of the systematic review. The dissertation category was more promising in terms of relevance. However, we could not access full texts for over 60% of the records in these categories. We experienced similar difficulties when trying to access full-text versions of book chapters (digital and/or hard copies). Considering all these factors, we decided to exclude 245 records of "grey" literature and book chapters, and to focus exclusively on published journal articles. These included articles from peer reviewed monographs or reports from reputable institutions with original data, as these papers were easily accessible and contained relevant and useful data.

We could not access the full texts of 43 out of the 804 journal articles (5%); these records were therefore excluded. The remaining 761 full-text articles were screened for relevance. Five hundred and sixty-five records (74%) were deemed irrelevant to the aims and scope of the systematic review. During the screening process, the reviewers discovered 13 possibly relevant journal articles that were not included in the search results. This resulted in a total of 209 journal articles with (possible) SSAn participants who completed the ASEBA forms. Of the 209 studies, 64 (31%) were excluded for reasons related to the sample description and use of the ASEBA forms. Specifically, 20 studies described the African sample vaguely (e.g., "participants from Africa", with no specific reference to country of origin), and a further 20 studies included SSAn participants in their samples but did not report stratified ASEBA data for those participants. Furthermore, 19 studies used or scored items in non-standardised ways. Examples of non-standardised use included administering a small selection of items independently (i.e., not as part of an established ASEBA subscale), using an incomplete subscale without justification, using a few items as part of a new measure, and altering the standard response format (e.g., from a three-point scale to a two-point scale). We did not have access to ASEBA forms published before 1991, so we could not ascertain whether the three studies that used these older versions administered full or partial subscales. Two studies (reporting data from the same sample) used a modified version of the YSR whose items bore little resemblance to those of the original YSR. Two studies administered the ASEBA forms as part of the study but did not report the actual data. Finally, we excluded

two case studies. A brief description of these excluded articles can be found in Table S2 in the digital supplement.

After these exclusions, we were left with 145 studies with specific data for SSAn participants that used the ASEBA forms in the standardised way. A few different articles stemmed from the same 'umbrella' study, as indicated by identical or overlapping samples. In addition, a few multi-nation studies utilised the same data set for secondary data analysis. Of the 145 studies, only 58 (40%) reported the psychometric properties of the ASEBA forms for the study sample and these studies were included in the final analysis. The digital supplement presents a list of the studies without reported psychometric properties (n = 87; Appendix S1).

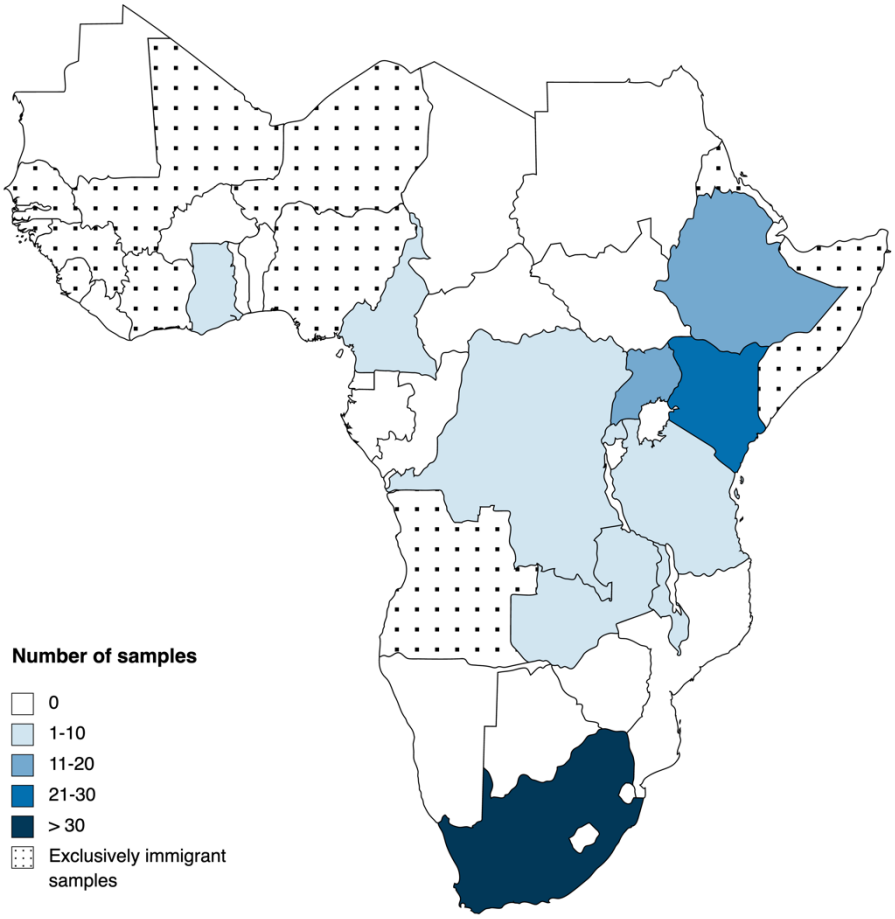
Study Characteristics

Region

Figure 4 illustrates the number of studies from the different SSAn countries. The map presents data from all 145 studies that met the first two inclusion criteria (i.e., SSAn participants and standardised use of the ASEBA forms), regardless of whether psychometric properties were reported. We displayed data from all 145 studies to visualise patterns of ASEBA usage across SSA. Six studies from outside of SSA comprised immigrant participants who originated from one of the following fourteen countries: Angola, Eritrea, Ethiopia, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Mali, Niger, Nigeria, Sierra-Leone, Senegal, and Somalia.

Figure 4

Map of Sub-Saharan Africa Displaying Number of ASEBA Studies Originating from Each Country (N = 145)



Note. ASEBA = Achenbach System of Empirically Based Assessment. The studies referenced in this map are those identified at the second stage of the review (i.e., sub-Saharan African sample and standard use of the ASEBA forms). The number of samples from each shaded country were as follows: South Africa (*n* = 70), Kenya (*n* = 27), Uganda (*n* = 19), Ethiopia (*n* = 11), Malawi (*n* = 4), Rwanda (*n* = 2), Tanzania (*n* = 2), Cameroon (*n* = 1), Democratic Republic of the Congo (*n* = 1), Ghana (*n* = 1), Zambia (*n* = 1). Immigrant participants' countries of origin from seven studies outside of sub-Saharan Africa are indicated with a dotted pattern: Angola, Eritrea, Gambia, Guinea, Guinea-Bissau, Mali, Niger, Nigeria, Sierra-Leone, Senegal, and Somalia. Immigrant samples also originated from Ethiopia, Ghana, and Kenya. Map created with mapchart.net.

An overview of the map shows the predominance of studies with samples from Southern African (n = 70, 48%) and East African (n = 66, 45%) countries. In Southern Africa, all samples originated from South Africa, while East African samples were distributed across a number of countries, including Kenya (n = 27, 41%), Uganda (n = 19, 29%), Ethiopia (n = 11, 17%), Malawi (n = 4, 6%), Rwanda (n = 2, 3%), Zambia (n = 1, 2%), and Tanzania (n = 1, 2%). In contrast, there were only two studies from Central Africa, which came from the Democratic Republic of the Congo (DRC) and Cameroon respectively, and only one study from West Africa, which came from Ghana. Many participants from immigrant samples originally came from countries in West Africa.

The distribution of studies across SSA was similar after narrowing the studies down to those that reported psychometric properties of the CBCL (N = 58). The breakdown of studies by region was as follows: Southern Africa (n = 29, 50%), East Africa (n = 25, 43%), Central Africa (n = 1, 2%) and immigrant samples (n = 3, 5%) from East and West Africa, living in Sweden (n = 1) and the USA (n = 2).

Relevant information about the studies including the reported psychometric properties (n = 58) is presented in Table 1.

Table 1

Description of Studies that Administered the ASEBA Forms to a Sub-Saharan African Sample and Reported their Psychometric Properties (N = 58)

Author(s)	Sample description	Version	Language(s) of administration	Translation/adaptation process	Purpose of the tool in the context of study aims	Psychometric properties
Southern Africa (*n* = 29)						
South Africa (*n* = 29)						
Allen et al. (2013)	HIV negative children (aged 6-10 years) with HIV positive mothers, living in a township community in Tshwane *N* = 361	CBCL/4-18 (1991)	Sepedi, isiZulu, Setswana, Sesotho, and English	• Translation • Back-translation • Reviewed by cultural advisors (community representatives) • Pilot test (*N* = 20)	Investigated relationships between maternal psychological functioning, parenting, and child behaviour in an HIV/AIDS context.	IC (Cronbach's α) Internalising Problems [α = 0.852] Externalising Problems [α = 0.915]
[a]Barber et al. (2005) *Linked to Bradford et al. (2003)*	Xhosa- (*n* = 600), Afrikaans- (*n* = 600), and English-speaking (*n* = 600) adolescents (aged 13-17 years) living in Cape Town	YSR (1987)	Xhosa, Afrikaans, and English	• Translation • Back-translation • Reviewed by community collaborators	Analysed links between parenting and adolescent functioning across countries/ethnic groups.	IC (Cronbach's α) *Delinquent Behaviour (8 items)* Xhosa [α = 0.75] Afrikaans [α = 0.79] English [α = 0.75]
Boyes et al. (2012); Boyes and Cluver (2013)	Children and adolescents, (aged 10-19 years) living in peri-urban communities in Cape Town *N* = 1025	YSR (1991)	Xhosa	• Translation • Back-translation • Reviewed by 5 health and social workers	Determined the psychometric properties of the Revised Children's Manifest Anxiety Scale and the Child PTSD Checklist.	IC (Cronbach's α) Somatic Complaints (9 items) [α = 0.66] Delinquent Behaviour [α = 0.61]
[a]Bradford et al. (2003)	Xhosa- (*n* = 635), Afrikaans- (*n* = 520), and English-speaking (*n* = 579) adolescents (aged 14-17 years) living in Cape Town	YSR (1987)	Xhosa, Afrikaans, and English	• Translation • Back-translation • Reviewed by community collaborators	Investigated the relationship between interparental conflict and child outcomes in nine countries.	IC (Cronbach's α) *Delinquent Behaviour (6 items)* Xhosa [α = 0.75] Afrikaans [α = 0.72] English [α = 0.70]

[#]Cluver, Lachman, et al. (2017)	Adolescents (aged 10-17 years) living in two poor rural communities in the Eastern Cape province *N* = 30	CBCL and YSR (1991)	Xhosa	• Translation • Back-translation	Evaluated the effectiveness of a parenting programme. Problem behaviour was measured pre- and post-intervention.	IC (Cronbach's α) *Rule-Breaking Behaviour (17 items)* CBCL [α = 0.82] YSR [α = 0.62] *Aggressive Behaviour (18 items)* CBCL [α = 0.85] YSR [α = 0.54]
Cluver, Meinck, Yakubovich, et al. (2016)	Adolescents (aged 10-17 years) living in poorly resourced rural and peri-urban communities. *N* = 115 (dyads)	CBCL/4-18 and YSR (1991)	Xhosa	• Translation • Back-translation	A pre-post trial testing the effectiveness of a parenting support programme to reduce child abuse.	IC (Cronbach's α) *Externalising Problems (35 items)* CBCL [α = 0.88] YSR [α = 0.71]
Collishaw et al. (2016)	Children and adolescents (aged 10-19) living in urban settlements in Cape Town Time 1: *N* = 1025 Time 2: *N* = 716	YSR (1991)	Xhosa	• Translation • Back-translation • Pilot test	Identified predictors of mental health resilience in children parentally bereaved by AIDS.	IC (Cronbach's α) *Delinquency (11 items)* Time 1 [α = 0.61] Time 2 [α = 0.64]
Cortina et al. (2013)	Primary school-going children (aged 10-12 years) from a poor rural area in Mpumalanga province *N* = 1025	YSR (1991)	Shangaan	• Translation • Back-translation • Pilot test (*N* = 200)	Estimated the prevalence of psychological problems amongst children in a poor rural school setting.	IC (Cronbach's α) Anxious/Depressed (13 items) [α = 0.63]
du Plessis et al. (2015)	Primary school children (aged 12-15 years) living in low-income communities in Cape Town *N* = 616	YSR (1983)	English and Afrikaans	• Translation • Back-translation • Pilot test (*N* = 72)	Examined associations between exposure to violence and internalising and externalising behaviours.	IC (Cronbach's α) *Delinquency (12 items)* Full sample [α = 0.70] English version [α = 0.67] Afrikaans translated version [α = 0.75]

Gwandure (2007)	School-going children (aged 14-19 years) providing home-based AIDS care for a parent/caregiver (*n* = 30) and a control group (*n* = 30)	CBCL/4-18 (1991)	NR	NR	Investigating the impact of providing home-based AIDS care on children's psychological and cognitive functioning.	IC (Cronbach's α) Total Problems [α = 0.87]
LeCroix et al. (2020) *Linked to Palin et al. (2009)*	Children (aged 11-16 years) with a mother living with HIV *N* = 104	CBCL (1991)	English, Afrikaans, and Sotho	• Translation • Back-translation • Reviewed by community collaborators • Pilot test and focus groups	Investigated the role of mother-child relationships in the context of maternal HIV and child outcomes.	CFA *Internalising Problems* The Somatic Complaints subscale was removed as the items did not load adequately. An additional 5 items were eliminated due to translation errors during the back-translation process. Of the remaining 18 items, 13 items loaded at 0.40 and above. *Externalising Problems* One item (i.e., sets fires) was excluded given feedback received from focus groups and piloting. Of the remaining 32 items, 23 items loaded at 0.40 and above. IC (Cronbach's α) Internalising Problems (13 items) [α = 0.77] Externalising Problems (23 items) [α = 0.87]
Malcolm-Smith et al. (2015)	Grade 1 learners (aged 6-8 years) from an English-medium school serving a predominantly lower middle-class community in Cape Town *N* = 65	CBCL (2001)	English	NA	Investigated associations between empathy and aggression in young school-going children.	IC (Cronbach's α) Externalising Problems [α = 0.87]

[#]Meinck et al. (2018) *Linked to Shenderovich et al. (2020)*	Adolescents (aged 10-18 years), living in the Eastern Cape Province *N* = 552	YSR (2001)	Xhosa	• Translation • Back-translation *Extracted from Cluver, Meinck, Shenderovich, et al. (2016)*	Investigated the psychometric properties of a child abuse self-report screening tool.	IC (Cronbach's α) Externalising Problems [α = 0.85]
[#]Meinck et al. (2017)	Adolescents (aged 10-17 years) from rural and urban districts in KwaZulu-Natal province *N* = 2477	YSR (2001)	IsiZulu	• Translation • Back-translation	Investigated the relationship between family-related factors (e.g., deprivation, parenting) and adolescent health outcomes.	IC (Cronbach's α) Delinquency [α = 0.64]
[#]Meinck et al. (2019)	Adolescent girls (aged 10 to 17 years) from urban and rural areas in two South African provinces *N* = 3516	YSR (2001)	Zulu, Xhosa, Sotho, Sepedi, Swati and Tsonga	• Translation • Back-translation	Investigated the moderating effects of free schooling in the relationship between adverse childhood experiences and HIV risk behaviours.	IC (Cronbach's α) Delinquency [α = 0.58]
Palin et al. (2009)	Children (aged 6-11 years) whose primary caregivers were woman living with HIV *N* = 103	CBCL/4-18 (1991)	Sotho, English, and Afrikaans	• Translation • Back-translation • Pilot test and focus groups	Studied the relationship between disclosure of maternal HIV infection and child behaviour.	CFA Somatic complaints items were eliminated for analyses, as factor loadings for these items were low (<0.40). An additional five items were eliminated due to translation errors. One item ("Sets fires") was excluded based on feedback from piloting and focus group. IC (Cronbach's α) Internalising Problems (13 items) [α = 0.78] Externalising Problems (23 items) [α = 0.88]

Peltzer and Pengpid (2013)	Children (aged 1.5-5 years) with a mother who had obtained a protection order against a male partner in Vhembe district, Limpopo province *N* = 86	CBCL/1.5-5 (2000)	Tsonga and Venda	• Translation • Back-translation	Examined associations between intimate partner violence and behavioural problems among pre-school children.	IC (Cronbach's α) Internalising Problems [α = 0.92] Externalising Problems [α = 0.94] Total Problems [α = 0.96]
#Rawatlal et al. (2015)	School-going adolescents (aged 9-18 years) in KwaZulu Natal province *N* = 206	CBCL/4-18 (1991)	English	NA	Examined the relationship between adolescent-parent attachment, perceived support from parents, and internalising symptoms.	IC (Cronbach's α) Depression/Anxiety (14 items) [α = 0.81]
#Rochat, Houle, et al. (2017)	HIV-exposed and unexposed children (aged 7-11 years) from Hlabisa sub-district in KwaZulu-Natal province *N* = 1536	CBCL/6-18 (2001)	Zulu	• Translation • Back-translation • Obtained translation license from ASEBA and translation was approved by test developer	Examined associations between early life factors, a breastfeeding intervention, and later child emotional and behavioural problems.	IC (Cronbach's α) Total Problems [α = 0.94]
#Rochat et al. (2015); #Rochat, Mitchell, et al. (2017)	HIV-uninfected children (aged 6-10 years) of HIV-infected women from the Hlabisa sub-district in KwaZulu Natal province *N* = 281	CBCL/6-18 (2001)	Zulu	• Translation • Back-translation • Obtained translation license from ASEBA and translation was approved by test developer	Examined child psychological outcomes of maternal HIV disclosure intervention and communication about HIV and death in a rural South African setting.	IC (Cronbach's α) *Total Problems* Pre-intervention [α = 0.94] Post-intervention [α = 0.92]
#Rochat et al. (2016)	HIV-negative children (aged 7-11 years) born to HIV-positive and HIV-negative mothers *N* = 906	CBCL/6-18 (2001)	Zulu	• Translation • Back-translation • Obtained translation license from ASEBA and translation was approved by test developer	Investigated the relationship between exclusive breastfeeding and child behavioural problems, in HIV-uninfected (exposed and unexposed) children.	IC (Cronbach's α) Total Problems [α = 0.94]

Shenderovich et al. (2020)	Adolescents (aged 10-18 years) living in the Eastern Cape *N* = 552	CBCL (2000)	Xhosa	• Translation • Back-translation *Extracted from Cluver, Meinck, Shenderovich, et al. (2016)*	Studied moderators of outcomes of a parenting intervention programme.	IC (Cronbach's α) Externalising Problems (35 items) [α = 0.90]
#Sipsma et al. (2013)	Young children (aged 6-10) with HIV-positive and HIV-negative mothers, living in Tshwane *N* = 509	CBCL/6-18 (2001)	Likely Sepedi, Setswana, Sesotho, isiZulu and English	• Translation • Back-translation • Reviewed by community advisors • Pilot test (*N* = 20)	Compared behaviour and psychological functioning of children with HIV-positive and HIV-negative mothers.	IC (Cronbach's α) Internalising Problems (α = 0.85) Externalising Problems (α = 0.92)
#Swain et al. (2017)	School-going adolescents (aged 9-18 years) living in a low socioeconomic community in KwaZulu Natal province *N* = 324	CBCL/4-18 (1991)	English	NA	Estimated the prevalence of posttraumatic stress in adolescents living in low-socioeconomic communities.	IC (Cronbach's α) Anxiety/Depression [α = 0.81] Somatic Complaints [α = 0.77]
#aThornton et al. (2019)	Children/adolescents (aged 11-18 years, 67% orphaned) *N* = 175	YSR (2001)	English	NA	Identified clinical risk factors amongst young people at risk of suicide, living in South Africa and Guyana.	IC (Cronbach's α) Attention Problems (10 items) [α = 0.58]; Anxious/Depressed (8 items) [α = 0.60)]; Withdrawn/Depressed (14 items) [α = 0.79]; Somatic Complaints (11 items) [α = 0.77]; Thought Problems (10 items) [α = 0.75]; Social Problems (11 items) [α = 0.73]
#van Westrhenen et al. (2019)	Children (aged 7-13 years) who experienced a traumatic event or abuse, living in and around Johannesburg *N* = 125	CBCL/4-18 (1991)	English	NA	Evaluated a creative arts psychotherapy group programme with children who experienced a trauma.	IC (Cronbach's α) *Total Problems* Pre-intervention (*n* = 125) [α = 0.96] Post-intervention (*n* = 37) [α = 0.96]

[#]Visser et al. (2018)	HIV-infected and HIV-uninfected children (aged 6-12 years) living in Tshwane *N* = 167	CBCL/6-18 (2001)	Sepedi, Setswana, Sesotho and isiZulu	• Translation • Back-translation • Reviewed by community advisors • Pilot test (*N* = 20)	Compared psychological problems of HIV-infected and HIV-uninfected children.	IC (Cronbach's α) Affective Problems (13 items) [α = 0.73]; Anxiety Problems (6 items) [α = 0.48]; Somatic Problems (7 items) [α = 0.57]; ADHD (7 items) [α = 0.74]; Oppositional Behaviour (5 items) [α = 0.69]; Conduct Problems (17 items) [α = 0.86]; Total DSM-oriented scales (55 items) [α = 0.91]
East Africa (*n* = 25)						
Kenya (*n* = 8)						
[a]Alampay et al. (2017)	Luo children (aged 7-10 years) *N* = 95	CBCL/4-18 (1991)	Luo	• Translation • Back-translation	Investigated whether justice and severity moderated the effects of corporal punishment on child internalising and externalising behaviours.	IC (Cronbach's α) *Internalising (31 items)* Mother-report (*n* = 94) [α = 0.794] Father report (*n* = 90) [α = 0.759] *Externalising (33 items)* Mother-report [α = 0.807] Father-report [α = 0.807]
[a]Gershoff et al. (2010)	Children (aged 6-10 years) from the Rachuonyo District of Nyanza province *N* = 17	CBCL and YSR (1991)	Dholuo	• Translation • Back-translation	Investigated the relationship between parenting discipline practices and child behaviours.	IC (Cronbach's α) *Aggression:* CBCL (20 items) [α = 0.73] YSR (19 items) [α = 0.81] *Anxiety/Depression:* CBCL (14 items) [α = 0.64] YSR (16 items) [α = 0.81]
*Harder et al. (2014)	Children/adolescents (aged 11-18 years) from a poor, informal settlement and surrounding areas in Nairobi *N* = 301	YSR (2001)	English and Kiswahili	• Translation • Back-translation • Approved by tool developer	Investigated the psychometric properties and factor structure of the YSR in a sample of Kenyan youth.	CFA, IC (Cronbach's α) *See Table 5.*

*Kariuki et al. (2016)	Children (aged 1–6 years) living in Kilifi *N* = 301	CBCL/1.5-5 (2000)	KiSwahili	• Obtained translation licence from ASEBA • Translation • Back-translation • Translations evaluated by 5 Kiswahili speakers • Focus groups and interviews (*N* = 90) • Pilot test (*N* = 50)	Investigated the psychometric properties of the pre-school version of the CBCL in a Kenyan sample	CFA Separate CFAs were conducted for each syndrome scale and for the two broad band scales. RMSEAs for all scales were < 0.08. Mean item factor loadings for syndrome scales ranged from 0.38 (Withdrawn) to 0.53 (Anxious/Depressed). IC (Cronbach's α) Total Problems (α = 0.95) Internalising Problems (α = 0.87) Externalising Problems (α = 0.86) Syndrome scales (α range = 0.53 – 0.86) TRR (Pearson's *r*) Test-retest reliability of Total Problems (*n* = 38) was significant *r* = 0.76, *p* < 0.0001, as was inter-informant agreement between mother and alternative caretaker (*n* = 17; *r* = 0.89, *p* < 0.001) and mother and father (*n* = 29; *r* = 0.81, *p* < 0.001).
Kumar et al. (2018)	Children/adolescents (aged 2-16 years) from two informal communities in Nairobi *N* = 394	CBCL/1.5-5 (2000) and CBCL/6-18 (2001)	Kiswahili (99%) and English	• Obtained permission from test developer • Translation • Back-translation	Examined the role of maternal adverse childhood experiences on child mental health	IC (Cronbach's α) Internalising Problems (α = 0.76 – 0.85) Externalising Problems (α = 0.87 – 0.89) *Different alpha statistics were calculated for the different age groups.*
*Magai et al. (2018)	Children/adolescents (aged 6-18 years) living in Kenya's Central Province *N* = 1022	CBCL/4-18 and YSR (1991)	Swahili	• Translated version obtained from ASEBA	Estimated the prevalence of emotional and behavioural problems in Kenyan youth.	CFA, IC (Cronbach's α) *See Tables 4 and 5.*

[#]Magai et al. (2021)	School-aged survivors of neonatal insults living in Kilifi, Kenya, and a comparison group (aged 6-12 years) *N* = 375	CBCL/6-18 (2001)	Kiswahili	• Translated version obtained from ASEBA	Examined the mental health and quality of life of child survivors of neonatal insults, and a control group.	IC (Cronbach's α) Alpha coefficients for Internalising Problems, Externalising Problems, and Total Problems ranged from 0.66-0.87
[#]Skinner et al. (2014)	Luo children (average age 8.46 years, *SD* = 0.64) from Kisumu *N* = 100	YSR (1991)	Dholuo	• Translation • Back-translation	Investigated childrearing violence and child adjustment after exposure to post-election violence in Kenya	IC (Cronbach's α) *Externalising Problems (30 items)* Time 1 [α = 0.72] Time 2 [α = 0.88] Time 3 [α = 0.75]
Kenya and Tanzania (*n* = 1)						
Dorsey et al. (2020)	Children (aged 7-13 years) who had experienced the death of a parent, living in urban and rural sites in Kenya (*n* = 230) and Tanzania (*n* = 230)	CBCL/6-18 (2001) and YSR (1991)	Kiswahili	• Translation • Back-translation	Tested the effectiveness of a trauma-focused cognitive behavioural therapy for children who experienced the death of a parent.	IC (Cronbach's α) Alpha coefficients for the CBCL and YSR Internalising Problems and Externalising Problems broadband scales ranged from 0.73-0.89) *The YSR was completed by children aged 11 years or older.*
Ethiopia (*n* = 7)						
[b]Betancourt et al. (2012)	Kunama refugees (aged 11-18 years) participating in an emergency education programme *N* = 153	YSR and CBCL (1991)	Kunamenga	• Translation • Back-translation • Discussions with community leaders and key informants • A few items were removed (e.g., items related to suicide and self-harm) based on community feedback	Assessed adolescent behavioural problems in the context of an education and psychosocial support programme for refugees.	IC (Cronbach's α) Internalising Problems (29 items) [α = 0.62] Externalising Problems (33 items) [α = 0.77] *Note. It is not clear whether the above statistics correspond to the CBCL or YSR.*

*Geibel et al. (2016)	Adolescents (aged 15-18 years) from Addis Ababa *N* = 134	YSR (2001)	Amharic	• Translation • Back-translation • Reviewed by expert panel • Pilot test	Evaluated the psychometric properties of the YSR as a mental health screening tool for young Ethiopians vulnerable to HIV.	IC (Cronbach's α), TRR, criterion-related validity (ROC) *See Table 5.*
*[b]Hall et al. (2014)	Somali children (mean age 11.02 years, *SD* = 2.90) living in refugee camps in Ethiopia *N* = 147	CBCL/6-18 and YSR (2001)	Somali	• Qualitative study to elicit local terms describing mental health problems • Translation • Back-translation • Sex-related items were removed based on feedback from the community	Established psychometric properties of instruments measuring mental health problems in Somali children and adolescents living in refugee camps.	IC (Cronbach's α), combined TRR and IRR (Pearson's *r*), criterion-related validity (ROC) *See Tables 4 and 5.*
[#]Isaksson et al. (2017)	Children (mean age = 10.3 years, *SD* = 0.43) of women living in the Butajira Rural Health Program (BRHP) area *N* = 358	CBCL/6-18 (2001)	Amharic	• Obtained an already-translated version • The authors made slight modifications to the translation to make it more suited for a rural community.	Investigated a possible relationship between maternal perinatal stressors and child emotional and behavioural functioning.	IC (Cronbach's α) Total Problems [α = 0.93]
*[a]Ivanova, Dobrean, et al. (2007)	Children/adolescents (aged 11-18 years) from a regional school sample (Mulatu, 1997) *N* = 677	CBCL/4-18 (1991)	Amharic	• Translation • Back-translation • Pilot test (*N* = 20) Extracted from Mulatu (1997)	Tested the eight-syndrome structure of the CBCL across 30 societies.	CFA *See Table 4.*
*[a]Ivanova, Achenbach, et al. (2007)	Children/adolescents (aged 11-18 years) from a regional school sample (Mulatu, 1997) *N* = 674	YSR (1991)	Amharic	• Translation • Back-translation • Pilot test (*N* = 20) Extracted from Mulatu (1997)	Tested the eight-syndrome structure of the YSR in 23 countries.	CFA *See Table 5.*

[b]Murray et al. (2018)	Children (mean age = 11.21 years, *SD* = 3.17), living in one of three refugee camps in the Somali region of Ethiopia *N* = 37	CBCL/6-18 and YSR (2001)	Somali	See Hall *et al*. (2014)	Evaluated a common elements treatment approach developed for mental illness/distress among children in Somali refugee camps.	IC (Cronbach's α) *CBCL* Internalising Problems [α = 0.95] Externalising Problems [α = 0.89] *YSR* Internalising Problems [α = 0.88] Externalising Problems [α = 0.93]
Uganda (*n* = 6)						
Amone-P'Olak et al. (2007)	Adolescents (aged 14-18 years) who were formerly abducted in Northern Uganda *N* = 294	YSR (1991)	NR	NR	Investigated the prevalence of war experiences and the use of cognitive emotion regulation in response to these experiences.	IC (Cronbach's α) Internalising Problems [α = 0.61] Externalising Problems [α = 0.61]
*Bangirana et al. (2009)	Child survivors of cerebral malaria (aged 7-16 years) *N* = 64	CBCL/6-18 (2001)	Luganda	• Translation • Back-translation	Investigated the psychometric properties of the CBCL in the context of a controlled trial aimed at improving cognitive functioning following cerebral malaria.	IC (Cronbach's α), test-rest reliability (Pearson's *r*) *See Table 4.*
Eggum et al. (2011)	Children (aged 9-20 years) residing near Tororo, Uganda *N* = 57	YSR (1991)	English	• Reviewed by expert panel for cultural appropriateness • Item "set fires" was removed	Determined prevalence of negative life effects and examined associations between negative life events and adjustment in Ugandan youth.	IC (Cronbach's α) Internalising Problems (20 items) [α = 0.75] Externalising Problems (27 items) [α = 0.83] *Some items were not administered because they were not culturally appropriate.*
[#]Familiar et al. (2015)	HIV-infected children (aged 5-12 years, *n* = 144) and children with a history of severe malaria (aged 5-12 years, *n* = 106)	CBCL/4-18 (1991)	Luganda	• Used a previously translated version from Bangirana *et al*. (2009)	Assessed structural overlap between the CBCL and the Behaviour Rating Inventory of Executive Function (BRIEF) in a sample of Ugandan children.	IC (Cronbach's α) *Total Problems* Malaria group [α = 0.93] HIV-infected group [α = 0.90]

Klasen et al. (2010)	Former child soldiers (aged 11-17 years) *N* = 330	YSR (2001)	Luo	• Translation • Back-translation • Reviewed by expert panel	Examined the effects of war and domestic violence on the mental health of child soldiers.	IC (Cronbach's α) Alpha coefficients ranged were between 0.70 and 0.95, except for Withdrawn/Depressed syndrome scale [α = 0.60] and Attention Problems syndrome scale [α = 0.64].
Ruiseñor-Escudero et al. (2015)	HIV-infected children (aged 5-12 years) living in Kayunga *N* = 144	CBCL/4-18 (1991)	Luganda	• Translation • Back-translation	Investigated the association between immunological parameters and behavioural problems among children living with HIV.	IC (Cronbach's α) Total Problems [α = 0.76]
Rwanda (*n* = 2)						
Ng et al. (2015)	Children (aged 10-17 years) living with HIV, children with a parent living with HIV, and children unaffected by HIV, living in Kayonza or Kirehe Districts *N* = 683	YSR (1991)	Kinyarwanda	• Translation	Investigated suicidal ideation and behaviours in children living with HIV compared to children with a parent living with HIV, and children unaffected by HIV.	IC (Cronbach's α) Internalising Problems (16 items) [α = 0.89]
Nsabimana et al. (2019)	Institutionalised and non-institutionalised children (aged 9-16 years) from rural and urban areas in Rwanda *N* = 178	CBCL/6-18 (2001)	Kinyarwanda	• Translation • Back-translation • Reviewed by expert panel	Explored the effects of institutionalization on externalising and internalising problems in children/adolescents.	IC (Cronbach's α) Internalising Problems [α = 0.84] Externalising Problems [α = 0.87]
Zambia (*n* = 1)						
*Murray et al. (2020)	Orphaned or otherwise vulnerable adolescents (aged 13-17 years) living in Lusaka *N* = 210	YSR (2001)	English and Nyanja	• Translation • Back-translation • Pilot test	Evaluated the psychometric properties of various measures for assessing psychological well-being.	CFA, IC (Cronbach's α), TRR (*ρ*), criterion validity (ROC), hypothesis testing for construct validity (*W*) *See Table 5.*

Central Africa (*n* = 1)						
Cameroon (*n* = 1)						
Wadji et al. (2020)	Children (aged 2-18 years) exposed to intimate partner violence *N* = 74	CBCL/4-18 (1991)	French and English	• Used an already-translated French version	Investigated the effects of maternal childhood abuse on child psychopathology.	IC (Cronbach's α) Internalising Problems [α = 0.84] Externalising Problems [α = 0.89]
Immigrant samples (*n* = 3)						
[a]Anakwenze & Rasmussen (2021); [a]Rasmussen et al. (2018)	Children (aged 5-12 years) with an immigrant parent from West Africa (Guinea, Sierra Leone, Senegal) who had been living in the USA for at least three years *N* = 91 (2021) *N* = 93 (2018)	CBCL/4-18 (1991)	Fulani and English	• Translation • Met with interviewers during data collection to address issues of translation and wording	(2021) Investigated the impact of parental trauma, parenting difficulty, and family separation on child behaviour. (2018) Investigated associations between migration and immigrant parents' perceptions of their children's safety.	IC (Cronbach's α) Externalising Problems (32 items) [α = 0.90]
[#]Osman et al. (2017)	Children (aged 11-16 years) with Somali-born parents, living in Sweden *N* = 109	CBCL/6-18 (2001)	Somali	• Obtained permission to translate the tool from tool developer • Translation • Back-translation • Pilot test	Evaluated a parenting support intervention for Somali-born parents and determined its effectiveness in reducing child behavioural problems.	IC (Cronbach's α) Cronbach's alpha coefficients for all syndrome scales were > 0.70. *Sex-related items (59, 60, 73, 96, and 110) were excluded from the questionnaire.*

Note. ASEBA = Achenbach System of Empirically Based Assessment, CBCL = Child Behaviour Checklist, YSR = Youth Self-Report, CFA = confirmatory factor analysis, IC = internal consistency, TRR = test-retest reliability, IRR = inter-rater reliability, ROC = receive operating characteristic, NR = not reported. Names of language and ethnic groups are reported verbatim from the articles. The same language is sometimes referred to by slightly different names (e.g., Xhosa = isiXhosa). The syndrome scale name 'Delinquent Behaviour/Delinquency' was updated to 'Rule-Breaking Behaviour' in the 2000/2001 versions. For journal articles corresponding to the same umbrella study: If the sample sizes and the psychometric statistics were markedly different, these articles were listed separately in different rows. If these differences were negligible, these articles are listed in a single row.

Articles marked with an asterisk (*) are psychometric studies (i.e., the primary focus of the study was the psychometric properties of the ASEBA forms).

[#]Corresponding author provided clarification and/or additional information upon request.

[a]Multi-nation studies.

[b]Sample comprises refugees from other sub-Saharan African countries.

ASEBA Versions

Studies were published between the years 2003 and 2021. Most studies used the most recent versions of the ASEBA forms published in 2000 and 2001 (*n* = 27, 47%) or older versions published in the 1990s (*n* = 29, 50%). One study used the 2001 CBCL/6-18 and the 1991 YSR. Three studies (5%) used versions published in the 1980s. The parent report CBCL forms were used more frequently (*n* = 38, 66%) than the YSR (*n* = 28, 48%). Eight studies (14%) administered both the CBCL and the YSR. No studies used the TRF.

Sample Characteristics

The majority of samples comprised school-aged children and adolescents, with only four studies (7%) including pre-school aged children. Participants across all regions were typically from poorly resourced communities or populations with specific vulnerabilities (e.g., refugees, orphans, survivors of trauma or violence, living with an HIV-infected parent, etc). Although most samples were drawn from communities or schools (i.e., non-clinical populations), six samples (10%) comprised children with one or more illnesses, the most common being HIV and cerebral malaria.

Sample sizes ranged from 17 (Gershoff et al., 2010) to 3516 (Meinck et al., 2019). The average sample size across all 58 studies was 493.16 (*SD* = 642.77) and the median sample size was 281 (*IQR* = 105.25 – 600).

Use of the ASEBA Forms

The CBCL and YSR were primarily used as an outcome measure of child behavioural and emotional problems. Two studies, one each from South Africa and Kenya, used the CBCL and YSR respectively to estimate the prevalence of child behavioural and emotional problems (Cortina et al., 2013; Magai et al., 2018). Twenty-nine studies (50%) used one or both of the broad band scales, ‘Internalising Problems’ and ‘Externalising Problems’, and/or all problem items as a single scale (‘Total Problems’, *n* = 16, 28%). Twenty-four studies (41%) used one or more of the syndrome scales individually (i.e., not as part of a broadband scale). Only four studies (7%) used one or more of the DSM-oriented scales.

Of the 58 studies, only nine (16%) were “psychometric” studies (i.e., where a primary focus of the study was the psychometric properties of the CBCL or YSR). Interestingly, although half of the studies were from South Africa, all psychometric studies came from East African countries, namely Ethiopia (*n* = 4), Kenya (*n* = 2), Uganda (*n* = 2), and Zambia (*n* = 1).

Languages of Administration, Translations, and Adaptations

Forty-nine studies (84%) used at least one translated version of the tool, while five studies (9%) administered the forms exclusively in English. Information regarding the language(s) of administration was not available for two studies. In South Africa, 10 out of 29 studies (34%) administered the tool in more than one language. Studies from other regions of SSA typically administered the ASEBA forms in one of the local languages (e.g., Swahili or Luo in Kenya). Two Kenya-based studies obtained official Swahili translations from ASEBA (Magai et al., 2018, 2021). At least one study from Uganda used the Luganda translation of the CBCL prepared by Bangirana and colleagues (2009). The study from Cameroon used an existing French translation of the CBCL/4-18 from another study (Wadji et al., 2020). One study from Ethiopia also used an existing Amharic translation of the CBCL/6-18 but slightly modified some translated items to improve their comprehensibility for a rural setting (Isaksson et al., 2017). Seven studies, one from Sweden (with Somali immigrants), two from Kenya, and four from the same umbrella study in South Africa, obtained a translation license from ASEBA.

The level of detail reported about the translation and adaptation processes varied considerably. Some studies included statements such as "all research materials were translated and back-translated". Others reported the translation process in great detail, including the number of people involved at each stage of the process, as well as each person's qualifications and areas of expertise. Published guidelines on how to approach translation and adaptation of tools vary somewhat but tend to have overlapping features (Sousa & Rojjanasrirat, 2011). According to COSMIN translation guidelines, a 'very good' translation process requires (i) at least two independent forward translators with a mother-tongue in the target language, one with expertise in the construct being measured, the other naïve on the construct being measured, (ii) at least two independent back-translators, naïve on the construct being measured, with a mother tongue in the source (original) language, (iii) a clear description of how discrepancies will be resolved, (iv) a review committee (excluding the translators, preferably including the tool developer), (v) a pilot study (e.g., cognitive interview) inspecting the content validity of the translated version with a sample representative of the target population, and (vi) a written feedback report on the translation process (Mokkink et al., 2019). In light of the inconsistent reporting of the translation processes in the included studies, we could not evaluate the translations using these guidelines.

All but three studies that conducted their own translations and adaptations (n = 44) reported using forward and back-translation methods. Thirteen studies reported using an expert panel – typically including cultural advisors, community representatives, local healthcare workers or mental health experts, and psychometricians – to evaluate the tool instructions, response format, and items for conceptual equivalence and cultural appropriateness. Three studies also conducted interviews and focus groups with members of the target population to rate the clarity of the instructions, response format, and individual items. Fourteen studies piloted the translated versions in samples ranging in size from 20 to 200 individuals. Most of these studies did not report detailed findings of the pilot testing or focus groups. However, a few studies removed some items based on community feedback. A study from Uganda did not administer "culturally inappropriate" items on the YSR, including "I set fires" (Eggum et al., 2011). Another two South African studies from the same umbrella study removed the item "sets fires" from the CBCL (LeCroix et al., 2020; Palin et al., 2009). Interestingly, one study each from Kenya and Ethiopia also removed this item post-hoc from the YSR and CBCL respectively, as it did not perform well in confirmatory factor analyses (Ivanova, Dobrean, et al., 2007; Magai et al., 2018). A study from Ethiopia removed items related to suicide and another two studies with Somali participants removed sex-related items (Hall et al., 2014; L.K. Murray et al., 2018; Osman et al., 2017).

Psychometric Properties

Fifty-six out of 58 studies (97%) reported internal consistency for one or more subscale using coefficient alpha (also known as Cronbach's alpha). There was substantial variation in the alpha coefficients reported for the same subscale. For example, alpha for the CBCL Internalising Problems scale across 16 studies ranged from 0.66 to 0.95 and for the same subscale on the YSR across nine studies from 0.61 to 0.95. There were too few studies in each country and language category to conduct a stratified reliability generalisation meta-analysis. Hence, we were not able to calculate an 'aggregated' internal consistency statistic for each translated version of the subscales.

Among the South African studies (n = 29), all but two reported only internal consistency statistic(s) for one or more subscales. Eleven studies from South Africa administered the tool in more than one language, but only three of those studies reported separate alpha statistics for the different translated versions. Two studies conducted separate CFAs for the Internalising Problems and Externalising Problems broad band scales but did not report detailed results for these analyses. All studies from Kenya (n = 10) reported coefficient alpha statistics, three conducted CFAs, and one study reported test-retest

reliability for the broad band scales, syndrome scales, and DSM-oriented scales. Studies from Ethiopia (*n* = 7) conducted more comprehensive psychometric analyses. Two studies (both by Ivanova et al., 2007 using data from Mulatu, 1997) conducted CFAs on the CBCL/4-18 and YSR respectively. Two studies evaluated criterion-related validity using receiver operating characteristic (ROC) curves. Finally, one investigated the test-retest reliability of the YSR, and another investigated combined test-retest and inter-rater reliability of both the CBCL/6-18 and the YSR. All studies from Uganda (*n* = 6) reported coefficient alpha and one study also reported test-retest reliability. The only study from Zambia conducted a comprehensive psychometric evaluation of the YSR, including CFA, internal consistency, criterion validity, test-retest reliability, and hypothesis testing. Studies from Tanzania (*n* = 1), Rwanda (*n* = 1), Cameroon (*n* = 1), and those from outside SSA with immigrant samples (*n* = 3) reported coefficient alpha only.

COSMIN Evaluation of the Psychometric-Focused Studies

We thoroughly reviewed eight of the nine psychometric-focused studies from East African countries using COSMIN guidelines. We excluded one study from Kenya, as it was the only study that used the CBCL/1.5-5 (Kariuki et al., 2016). Four of the eight studies evaluated the School-Age version of the CBCL, and six evaluated the YSR (two studies administered both). We decided to limit the COSMIN evaluation to psychometric-focused studies only, as the remaining studies (*n* = 49) reported only internal consistency statistics, typically for one or two subscales.

Table 2 displays the COSMIN criteria for good measurement properties, with a visual icon allocated to each rating. We added a fourth rating, ‘mixed results’, to indicate a measurement property with different ratings for different sub-groups of participants. The table lists only the measurement properties and criteria that were relevant to the included studies. In this review, there were no published studies that specifically evaluated the content validity, cross-cultural validity, measurement error, or responsiveness of the ASEBA forms.

Table 2
Adapted COSMIN Criteria for Adequacy of Measurement Properties

Rating	Symbol	Measurement property				
		Structural validity	**Internal consistency**	**Reliability**	**Criterion validity**	**Hypothesis testing**
Sufficient	😀	CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08	At least low evidence for sufficient structural validity AND Cronbach's alpha OR omega ≥ 0.70 for each unidimensional scale or subscale	Correlation between scores at different time points ≥ 0.70	Correlation with gold standard ≥ 0.70 OR AUC ≥ 0.70	The result is in accordance with the hypothesis (at least 75% of the results are in accordance with the hypotheses); correlations with instruments measuring similar constructs should be greater than 0.50, correlations with instruments measuring related, but dissimilar constructs should be lower (i.e., 0.30-0.50), correlations with instruments measuring unrelated constructs should be lower than 0.30, and there should be meaningful changes between relevant (sub)groups
Indeterminate	?	Not all information for 'adequate' reported	Criteria for "at least low evidence for sufficient structural validity" not met		Not all information for 'sufficient' reported	No hypothesis defined (by the review team)
Insufficient	☹	Criteria for 'adequate' not met	At least low evidence for sufficient structural validity AND Cronbach's alpha(s) or omega < 0.70 for each unidimensional scale or subscale	Correlation between scores at different time points < 0.70	Correlation with gold standard < 0.70 OR AUC < 0.70	The result is not in accordance with the hypothesis (75% of the results are not in accordance with the hypotheses)
Mixed results	⚖					

Note. CFA = confirmatory factor analysis, AUC = area under the curve, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Residuals.

We relaxed three standards set out by COSMIN. First, the guidelines state that reliability for ordinal scales should be estimated using weighted Cohen's Kappa (*k*). However, as no studies reported this statistic, we accepted Pearson's or Spearman's correlation coefficients as an acceptable method to estimate reliability. Second, with regards to hypothesis testing for construct validity, COSMIN recommends that reviewers formulate a set of hypotheses about the expected magnitude and direction of the correlations between measures and mean differences in scores between groups, based on theoretical understandings of the construct, prior to the review. This is intended to reduce the possibility of bias and to ensure standardisation across studies. This was not feasible for the present review as child behaviour is such a broad construct with many possible correlates. Hence, we accepted hypotheses as long as the authors provided evidence to substantiate them. Third, for the internal consistency standard, COSMIN also requires "at least low evidence for sufficient structural validity". However, we removed this standard for individual studies as the ASEBA forms are well-established worldwide and many studies have confirmed their structural validity (Ivanova, Achenbach, et al., 2007; Ivanova, Dobrean, et al., 2007). COSMIN also recommends that reviewers determine a reasonable "gold standard" prior to assessing the methodological quality of criterion-validity studies. We decided that a clinical assessment by a qualified mental health professional (e.g., registered clinical psychologist, psychiatrist, social worker) based on a standardised and validated diagnostic tool (e.g., DSM-5), could be considered a gold standard in this context. Table 3 displays the COSMIN criteria for evaluating the methodological quality of studies reporting psychometric properties. We assigned a colour to each rating for ease of reading. Table 4 and Table 5 present the combined results of the COSMIN analysis.

Table 3
COSMIN Risk of Bias Checklist for Evaluating Methodological Standards of Measurement Properties

Rating	Measurement property				
	Structural validity	**Internal consistency**	**Reliability**	**Criterion validity[b]**	**Hypothesis testing**
Very good	CFA performed; *N* is 7 times the number of items and ≥100; no other important methodological flaws (e.g., inappropriate estimation method)	Internal consistency statistic calculated for each unidimensional scale or subscale; Cronbach's alpha or omega calculated	Evidence provided that participants were stable in the interim period on the construct being measured; time interval appropriate[a]; test conditions were similar (evidence provided); Pearson/Spearman correlation coefficient calculated; no other important methodological flaws	AUC calculated OR sensitivity and specificity calculated; no other important methodological flaws	Constructs measured by comparator instrument is clear; sufficient measurement properties of comparator instrument in a population similar to study population OR adequate description of the important characteristics of the subgroups; statistical method applied was appropriate
Adequate	EFA performed; *N* is at least 5 times the number of items and ≥100 OR at least 6 times number of items but <100		Assumable that patients were stable; assumable that test conditions were similar		Sufficient measurement properties of comparator instrument but unsure if these apply to study population OR adequate description of most of the important characteristics of the subgroups; assumable that statistical method was appropriate
Doubtful	*N* is 5 times the number of items, but <100; Other minor methodological flaws	Unclear whether scale or subscale is unidimensional; only item-total correlations calculated	Unclear if patients were stable; doubtful whether time interval was appropriate or time interval was not stated; unclear if test conditions were similar; other minor methodological flaws	Other minor methodological flaws	Some information on measurement properties of the comparator instrument in any study population OR poor/no description of the important characteristics of the subgroups; statistical method applied NOT optimal
Inadequate	No EFA or CFA performed; *N* is < 5 times the number of items; Other important methodological flaws	Internal consistency statistic NOT calculated for each unidimensional scale or subscale; no Cronbach's alpha / omega and no item-total correlations calculated	Patients were NOT stable; test conditions were NOT similar; other important methodological flaws	AUC NOT calculated OR sensitivity and specificity NOT calculated; other important methodological flaws	Constructs measured by comparator instrument unclear; no information on the measurement properties of the comparator instrument(s) OR evidence for insufficient measurement properties of the comparator instrument; statistical method applied NOT appropriate

Note. CFA = confirmatory factor analysis, EFA = exploratory factor analysis, AUC = area under the curve.
[a]COSMIN recommends an interval period of two weeks.
[b]We decided that a clinical assessment by a qualified mental health professional (e.g., registered clinical psychologist, psychiatrist, social worker) based on a standardised and validated diagnostic tool (e.g., DSM-5), could be considered a gold standard in this context.

Table 4

Measurement Properties of the CBCL/6-18 and Risk of Bias Analysis from Four Psychometric Studies Based on COSMIN Guidelines

Measurement Property	*N*	Country	Broad band scales			Syndrome Scales								DSM-Oriented Scales					
			TOP	IP	EP	AB	AD	AP	RB	SP	SC	TP	WD	AFP	ANP	SOP	ADP	ODP	COP
Structural validity																			
Ivanova, Dobrean, et al. (2007)	677	Ethiopia				😀	😀	😀	😀	😀	😀	😀	😀						
Magai et al. (2018)	1022	Kenya				😀	😀	😀	😀	😀	😀	😀	😀						
Internal consistency																			
Bangirana et al. (2009)	64	Uganda	😀	☹	😀	😀	☹	☹	☹	☹	☹	☹	☹	☹	☹	☹	☹	☹	☹
Hall et al. (2014)	147	Ethiopia		😀	😀	😀	😀	😀	😀	😀	😀	😀	☹						
Magai et al. (2018)	1022	Kenya	😀	😀	😀	😀	⚖	😀	☹	☹	😀	⚖	☹						
Reliability																			
Bangirana et al. (2009)	22	Uganda	☹	☹	😀	😀	☹	☹	☹	😀	😀	☹	😀	☹	☹	☹	☹	😀	☹
Hall et al. (2014)	17	Ethiopia		☹	☹	☹	☹	☹	☹	☹	☹	😀	☹						
Criterion validity																			
Hall et al. (2014)	79	Ethiopia		😀															
Hypothesis testing																			
Hall et al. (2014)	174	Ethiopia		😀		☹	😀	?	☹	?	😀	?	😀						

Note. CBCL/6-18 = Child Behaviour Checklist for Ages 6-18. TOP = Total Problems, IP = Internalising Problems, EP = Externalising Problems, AB = Aggressive Behaviour, AD = Anxious/Depressed, AP = Attention Problems, RB = Rule-breaking Behaviour, SP = Social Problems, SC = Somatic Complaints, TP = Thought Problems, WD = Withdrawn/Depressed, AFP = Affective Problems, ANP = Anxiety Problems, SOP = Somatic Problems, ADP = Attention-Deficit/Hyperactive Problems, ODP = Oppositional Defiant Problems, COP = Conduct Problems. For more information about the Risk of Bias ratings for individual studies, please contact the corresponding author.

Table 5

Measurement Properties of the YSR and Risk of Bias Analysis from Six Psychometric Studies Based on COSMIN Guidelines

	N	Country	Broad band scales			Syndrome Scales								DSM-Oriented Scales					
			TOP	IP	EP	AB	AD	AP	RB	SP	SC	TP	WD	AFP	ANP	SOP	ADP	ODP	COP
Structural validity																			
Harder et al. (2014)	301	Kenya				😀	😀	😀	😀	😀	😀	😀	😀						
Ivanova, Achenbach, et al. (2007)	674	Ethiopia				😀	😀	😀	😀	😀	😀	😀	😀						
Magai et al. (2018)	533	Kenya				😀	😀	😀	😀	😀	😀	😀	😀						
Murray et al. (2020)	210	Zambia		😀	😀														
Internal consistency																			
Geibel et al. (2016)	134	Ethiopia		😀	⚖	☹	⚖	☹	☹	☹	☹	☹	☹	⚖	☹	☹	☹	⚖	☹
Hall et al. (2014)	147	Ethiopia		😀	😀	😀	😀	😀	😀	😀	😀	😀	😀						
Harder et al. (2014)	242	Kenya	😀	😀	😀	😀	☹	😀	☹	☹	😀	☹	☹						
Magai et al. (2018)	1022	Kenya	😀	😀	😀	☹	☹	☹	☹	☹	😀	☹	☹						
Murray et al. (2020)	210	Zambia		😀	😀														
Reliability																			
Geibel et al. (2016)	111	Ethiopia		⚖	☹	☹	☹	☹	☹	☹	☹	☹	☹	⚖	☹	☹	☹	☹	☹
Hall et al. (2014)	17	Ethiopia		☹	☹	☹	☹	☹	☹	☹	☹	😀	☹						
Murray et al. (2020)	33	Zambia		☹	☹														
Criterion validity																			
Geibel et al. (2016)	134	Ethiopia				☹	⚖	😀	☹	⚖	☹		☹						
Murray et al. (2020)	165	Zambia		😀	😀														
Hypothesis testing																			
Geibel et al. (2016)	134	Ethiopia				☹	⚖	☹	☹	⚖	☹		☹						
Hall et al. (2014)	174	Ethiopia				☹	😀	❓	☹	❓	😀	❓	😀						
Harder et al. (2014)	301	Kenya	😀	😀	😀	😀	⚖	😀	😀	😀	😀	😀	☹						
Harder et al. (2014)	242	Kenya		😀	☹	☹	😀		☹		😀		😀	😀	😀	😀			
Murray et al. (2020)	165	Zambia		😀	😀														
Murray et al. (2020)	210	Zambia		☹	☹														

Note. YSR = Youth Self Report. TOP = Total Problems, IP = Internalising Problems, EP = Externalising Problems, AB = Aggressive Behaviour, AD = Anxious/Depressed, AP = Attention Problems, RB = Rule-breaking, SP = Social Problems, SC = Somatic Complaints, TP = Thought Problems, WD = Withdrawn/Depressed, AFP = Affective Problems, ANP = Anxiety Problems, SOP = Somatic Problems, ADP = Attention-Deficit/Hyperactive Problems, ODP = Oppositional Defiant Problems, COP = Conduct Problems. For more information about the Risk of Bias ratings for individual studies, please contact the corresponding author.

All studies measuring structural validity conducted a CFA using tetrachoric or polychoric correlation matrices with a robust weighted least squares estimator, which is recommended for ordinal data (Li, 2016). In terms of structural validity, two 'very good' quality studies, one each from Kenya and Ethiopia, supported the factorial structure of the CBCL syndrome scales using CFAs. The YSR syndrome scales also performed well in CFAs across all four studies, although the methodological qualities of three studies were somewhat compromised by the smaller sample sizes. These findings suggest that the latent constructs measured by the ASEBA syndrome scales are being adequately explained by the specific behavioural problems (i.e., items) in these populations (De Kock et al., 2013).

Internal consistency was the most commonly reported measurement property, with three and five studies reporting coefficient alpha for CBCL and YSR subscales respectively. The COSMIN methodological standards for internal consistency are minimal, hence the overall quality of the methods used for this measurement property were 'very good' overall. Alpha coefficients for the broadband scales were generally higher than those for the syndrome scales and DSM-oriented scales. This was not surprising as the value of alpha is influenced by the length of the tool (Cortina, 1993). For the CBCL, the Aggressive Behaviour, Attention Problems, and Somatic Complaints syndrome scales performed the best in terms of internal consistency, meeting the 'sufficient' criteria in at least two out of the three studies. The Withdrawn/Depressed syndrome scale, however, did not meet the necessary criteria in any of the three studies. One study in this category had a relatively smaller sample size (n = 64; Bangirana et al., 2009) than the other two studies measuring internal consistency of the CBCL. In this smaller study, all but one of the syndrome and DSM-oriented scales did not meet the 'sufficient' criteria. For the YSR, Somatic Complaints was the only syndrome scale to meet the criteria in at least three of the four studies. Only two studies (one each for the CBCL and YSR) estimated internal consistencies for the DSM-oriented subscales, and these results were consistently insufficient.

The methodological quality of the reliability analyses for both the CBCL and the YSR were either 'doubtful' or 'inadequate'. Reasons for the poorer quality of studies were too long or too short time intervals between administrations (e.g., 9 weeks, 5-7 days), participants undergoing an intervention in between administrations, and other methodological flaws, including not specifying how a subset of the sample was selected for re-administration, and a lack of evidence that participants were stable on the construct to be measured in between administrations. In terms of the measurement property itself, the correlation coefficients were

consistently insufficient across forms and subscales. Overall, current evidence for test-retest reliability of the ASEBA forms is inadequate in these SSAn populations.

In terms of criterion validity, Geibel and colleagues' (2016) study was the only one to use a psychiatric assessment as a gold standard for the ROC analysis. However, these assessments were not based on a standardized clinical diagnostic tool. Two studies developed their own criteria to identify cases (Hall et al., 2014; S.M. Murray et al., 2020). Murray and colleagues (2020) created a four-item screening questionnaire, administered to the child and their caregiver, asking whether the child had significant psychosocial problems ("yes" or "no"). In Hall and colleagues' (2014) study, refugee camp social workers identified cases using a list of common internalising and externalising symptoms. The social workers' assessments were then corroborated with caregivers' responses to a short screening questionnaire. Based on the few studies included in this analysis, there is very limited evidence to substantiate the criterion validity of the ASEBA forms in SSA.

Four studies conducted some form of hypothesis testing to examine the construct validity of the ASEBA subscales. All four studies evaluated known-groups validity based on clinical characteristics of the sample (i.e., 'case' vs 'non-case') with mixed results, and one study from Kenya examined sex differences in levels of internalising and externalising behaviours respectively (Harder et al., 2014). The study from Zambia estimated convergent and divergent validity of the Internalising Problems and Externalising Problems subscales on the YSR (S.M. Murray et al., 2020). Only two out of the five comparator measures used (measuring post-traumatic stress and well-being, respectively) had adequate psychometric properties in a similar population. For both the Internalising Problems and Externalising Problems subscales, less than 75% of the results were in accordance with the hypotheses.

Discussion

We identified 145 studies that used the ASEBA forms to measure child behaviour problems in SSAn samples. This suggests that the ASEBA forms are used frequently, at least for research purposes, in SSAn contexts. However, less than half of the studies reported any measurement properties of the ASEBA forms. Of the studies that did report measurement properties, most reported only coefficient alpha as a measure of internal consistency for the subscales used. The widespread use of the ASEBA forms in SSA without evaluation of measurement properties warrants consideration. A tool's measurement properties are inextricably tied to the context in which it is administered. Without sufficient evidence to support the validity of the information derived from the tools used, the dependability of

results remains questionable. The tendency of applied researchers to conduct and report limited psychometric evaluations only (i.e., coefficient alpha), without any further investigation or interpretation, remains a challenge to research in this field. Comprehensive psychometric analyses are necessary to arrive at meaningful and accurate conclusions about a tool's measurement properties (Dima, 2018). In addition, psychometric analyses should be reported clearly and comprehensively, and this information should be easily accessible to readers. COSMIN is in the process of developing a checklist for standards on reporting measurement properties (see https://www.cosmin.nl/tools/checklists-assessing-methodological-study-qualities/). This will hopefully aid in developing a standardised and transparent approach to reporting measurement properties in research studies.

Most of the studies included in the final analysis administered at least one translated version of the CBCL or YSR, and almost all translations were created specifically for use in those studies. There were inconsistencies with regards to the reporting of the translation procedures. Descriptions of the translation procedures provided little detail, raising doubts about the integrity of the translations, as judged by COSMIN standards. Although it is possible that rigorous methodological guidelines were adhered to, this information was not readily available in most cases. Consequently, we were not able to evaluate the translation procedures across studies. The quality of a translation may significantly impact the validity of a tool (De Kock et al., 2013; Van Widenfelt et al., 2005). In a sense, a translated version of a tool becomes its own outcome measure that should be evaluated for content validity (Terwee, Prinsen, Chiarotto, de Vet, et al., 2018). Although a few studies reported the use of focus groups and pilot testing to assess relevance and comprehensibility, the results of these investigations (e.g., any re-wordings or modifications made to the original draft) were not always reported. Transparent reporting of translations and adaptations serves two important purposes. First, it grants readers the opportunity to evaluate the validity of the translated versions. Second, it serves as a useful record for researchers or clinicians who may be interested in administering the translated tool in future studies or in clinical settings. In this review, we found that only three out of the ten studies that administered the ASEBA forms in more than one language conducted separate psychometric analyses for each version. This would be considered an important step to rule out potential measurement bias.

In this review, we also evaluated eight of the nine psychometric-focused studies using COSMIN standards and criteria for good measurement properties. To our knowledge, these nine studies are the only published journal articles addressing the validity of the ASEBA forms in SSAn contexts. Overall, evidence to support the validity and reliability of the CBCL

and YSR in SSAn countries in the existing literature is limited. Furthermore, the variable quality of the methods used across different studies to assess the measurement properties of the CBCL and YSR preclude us from making confident recommendations regarding its use in these regions.

Having said this, the statistical methods used, as assessed by the COSMIN Risk of Bias Checklist, were generally adequate. The main exceptions to this were the reliability and criterion validity analyses. More studies with different designs and larger samples are needed to learn about the criterion validity, test-retest reliability, and inter-rater reliability of the ASEBA forms in SSA. Criterion validity is a very important measurement property if the ASEBA forms are to be used as screening tools in community settings. Although no single "gold-standard" instrument currently exists for child behavioural and emotional problems, judging ASEBA scores against clinical assessments based on standardized diagnostic tools may be a strong starting point.

Coefficient alpha was the most frequently reported statistic across all studies. However, there are limitations of coefficient alpha as a measure of internal consistency (Dunn et al., 2014; Sijtsma, 2009). Ordinal coefficient alpha may generate a more reliable estimate of internal consistency for Likert scales, such as the ASEBA forms (Zumbo et al., 2007). In terms of structural validity, the majority of studies were of a very good standard, barring a few studies with sample sizes smaller than required for a tool measuring multiple constructs with many items. Although the broadband scales (Internalising Problems and Externalising Problems) were frequently administered in SSA, no studies conducted higher order or bifactor CFAs that would have investigated the unidimensionality of these broad band scales.

Most of the studies included in the final analysis came from South Africa, although none of these studies were specifically focused on the measurement properties of the ASEBA forms. Hence, there remains limited evidence to support the validity of the ASEBA forms in a South African context. A smaller but significant proportion of the included studies came from East African countries (notably Kenya, Ethiopia, and Uganda). All nine psychometric-focused studies came from East African countries. Compared to Southern and East Africa, there were very few studies from West and Central Africa. It is possible that other measures of child and behavioural problems are more popular in these regions. Although there was only one study that came from a West African country (i.e., Ghana), individuals of West African origin living outside of SSA were also represented in the included studies.

Limitations

The two reviewers made every effort to ensure that all papers were thoroughly screened and reviewed. However, it is possible that a few articles were either not included in the search results, accidentally removed from the reference library, or incorrectly screened. A limitation of our study was the exclusion of unpublished "grey" literature, including theses, books, and conference presentations. Although we made the decision to exclude these records for practical reasons, grey literature would have likely enriched our analysis and reduced the risk of publication bias. Another important limitation of our study was that we could not use the COSMIN GRADE approach to quantitatively pool the results from individual studies and grade the overall quality of evidence for each measurement property (Mokkink, Prinsen, et al., 2018). Results from individual studies were too inconsistent to pool quantitatively. Moreover, there were too few studies in each "sub-group" (e.g., country, language of administration, sample characteristics) to arrive at reliable conclusions for each possible combination of subscale and sub-group.

Gaps Identified and Recommendations

One important gap in the current literature is the dearth of studies evaluating the content validity of the ASEBA forms in SSA. Content validity, the extent to which the content of a tool adequately represents the construct it measures, is arguably the most important of all measurement properties (Terwee, Prinsen, Chiarotto, Westerman, et al., 2018). If a tool does not have content validity, then all other measurement properties are irrelevant. As described earlier, there were some attempts to evaluate the relevance and comprehensibility of the ASEBA forms through pilot testing. To our knowledge, only one included study explored the comprehensiveness of the ASEBA forms. Prior to conducting their study, Hall and colleagues (2014) used qualitative methods to identify local symptoms of internalising and externalising behaviours in Somali refugees living in Ethiopia. The authors added 11 and 4 of these locally derived symptoms to the Internalising Problems and Externalising Problems subscales respectively. Although we could not include this preliminary study in the current analysis, the findings emphasise the importance of evaluating the comprehensiveness of behavioural screening tools in SSA.

Summary and Conclusion

The primary aim of the present review was to investigate the measurement properties of the ASEBA forms in SSAn countries, where translated versions of the forms are frequently administered. At present, evidence is limited in terms of both the number and quality of available studies. East African countries have already made significant progress with regards

to evaluating translated versions of the ASEBA forms in local contexts. In South Africa, however, the measurement properties of the ASEBA forms remain under-studied despite their widespread use in research. Data from other areas of SSA are largely absent. This review has demonstrated the importance of validating existing behavioural tools for culturally and linguistically diverse contexts in SSA. Comprehensive and ongoing psychometric evaluations of tools require time and resources. However, the result is that clinicians and researchers become more confident that the inferences made based on these tools are accurate and dependable.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. University of Vermont.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. University of Vermont.

Alampay, L. P., Godwin, J., Lansford, J. E., Bombi, A. S., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Tirado, L. M. U., Zelli, A., Al-Hassan, S. M., & Bacchini, D. (2017). Severity and justness do not moderate the relation between corporal punishment and negative child outcomes: A multicultural and longitudinal study. *International Journal of Behavioral Development*, *41*(4), 491–502. https://doi.org/10.1177/0165025417697852

Allen, A. B., Finestone, M., Eloff, I., Sipsma, H., Makin, J., Triplett, K., Ebersöhn, L., Sikkema, K., Briggs-Gowan, M., Visser, M., Ferreira, R., & Forsyth, B. W. C. (2013). 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*(3), 605–616. https://doi.org/10.1007/s10461-013-0544-7

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

Amone-P'Olak, K., Garnefski, N., & Kraaij, V. (2007). Adolescents caught between fires: Cognitive emotion regulation in response to war experiences in Northern Uganda. *Journal of Adolescence*, *30*(4), 655–669. https://doi.org/https://doi.org/10.1016/j.adolescence.2006.05.004

Anakwenze, O., & Rasmussen, A. (2021). The impact of parental trauma, parenting difficulty, and planned family separation on the behavioral health of West African immigrant children in New York City. *Psychological Trauma: Theory, Research, Practice, and Policy*. Advance online publication. https://doi.org/10.1037/tra0001011

Atilola, O. (2015). Level of community mental health literacy in sub-Saharan Africa: Current studies are limited in number, scope, spread, and cognizance of cultural nuances. *Nordic Journal of Psychiatry*, *69*(2), 93–101. https://doi.org/10.3109/08039488.2014.947319

Bangirana, P., Nakasujja, N., Giordani, B., Opoka, R. O., John, C. C., & Boivin, M. J. (2009). Reliability of the Luganda version of the Child Behaviour Checklist in measuring behavioural problems after cerebral malaria. *Child and Adolescent Psychiatry and Mental Health*, *3*, Article 38. https://doi.org/10.1186/1753-2000-3-38

Barber, B. K., Stolz, H. E., & Olsen, J. A. (2005). Parental support, psychological control, and behavioral control: Assessing relevance across time, culture, and method. *Monographs of the Society for Research in Child Development*, *70*(4), i–147.

Betancourt, T. S., Yudron, M., Wheaton, W., & Smith-Fawzi, M. C. (2012). Caregiver and adolescent mental health in Ethiopian Kunama refugees participating in an emergency education program. *Journal of Adolescent Health*, *51*(4), 357–365. https://doi.org/10.1016/j.jadohealth.2012.01.001

Boyes, M. E., & Cluver, L. D. (2013). Performance of the Revised Children's Manifest Anxiety Scale in a sample of children and adolescents from poor urban communities in Cape Town. *European Journal of Psychological Assessment*, *29*(2), 113–120. https://doi.org/10.1027/1015-5759/a000134

Boyes, M. E., Cluver, L. D., & Gardner, F. (2012). Psychometric properties of the Child PTSD Checklist in a community sample of South African children and adolescents. *PLoS One*, *7*(10), Article e46905. https://doi.org/10.1371/journal.pone.0046905

Bradford, K., Barber, B. K., Olsen, J. A., Maughan, S. L., Erickson, L. D., Ward, D., & Stolz, H. E. (2003). A multi-national study of interparental conflict, parenting, and adolescent functioning: South Africa, Bangladesh, China, India, Bosnia, Germany, Palestine, Colombia, and the United States. *Marriage & Family Review*, *35*(3–4), 107–137. https://doi.org/10.1300/J002v35n03_07

Cluver, L. D., Lachman, J. M., Ward, C. L., Gardner, F., Peterson, T., Hutchings, J. M., Mikton, C., Meinck, F., Tsoanyane, S., Doubt, J., Boyes, M., & Redfern, A. A. (2017). Development of a parenting support program to prevent abuse of adolescents in South Africa: Findings from a pilot pre-post study. *Research on Social Work Practice*, *27*(7), 758–766. https://doi.org/10.1177/1049731516628647

Cluver, L., Meinck, F., Shenderovich, Y., Ward, C. L., Romero, R. H., Redfern, A., Lombard, C., Doubt, J., Steinert, J., Catanho, R., Wittesaele, C., De Stone, S., Salah, N., Mpimpilashe, P., Lachman, J., Loening, H., Gardner, F., Blanc, D., Nocuza, M., & Lechowicz, M. (2016). A parenting programme to prevent abuse of adolescents in South Africa: Study protocol for a randomised controlled trial. *Trials*, *17*, Article 328. https://doi.org/10.1186/s13063-016-1452-8

Cluver, L., Meinck, F., Yakubovich, A., Doubt, J., Redfern, A., Ward, C., Salah, N., De Stone, S., Petersen, T., Mpimpilashe, P., Romero, R. H., Ncobo, L., Lachman, J., Tsoanyane, S., Shenderovich, Y., Loening, H., Byrne, J., Sherr, L., Kaplan, L., & Gardner, F. (2016). Reducing child abuse amongst adolescents in low- and middle-income countries: A pre-post trial in South Africa. *BMC Public Health*, *16*, Article 567. https://doi.org/10.1186/s12889-016-3262-z

Collishaw, S., Gardner, F., Aber, J. L., & Cluver, L. (2016). Predictors of mental health resilience in children who have been parentally bereaved by AIDS in urban South Africa. *Journal of Abnormal Child Psychology*, *44*(4), 719–730. https://doi.org/10.1007/s10802-015-0068-x

Cortina, J. M. (1993). What Is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Cortina, M. A., Fazel, M., Hlungwani, T. M., Kahn, K., Tollman, S., Cortina-Borja, M., & Stein, A. (2013). Childhood psychological problems in school settings in rural Southern Africa. *PLoS One*, *8*(6), Article e65041. https://doi.org/https://doi.org/10.1371/journal.pone.0065041

Cortina, M. A., Sodha, A., Fazel, M., & Ramchandani, P. G. (2012). Prevalence of child mental health problems in Sub-Saharan Africa: A systematic review. *Archives of Pediatrics and Adolescent Medicine*, *166*(3), 276–281. https://doi.org/10.1001/archpediatrics.2011.592

De Kock, F., Kanjee, A., & Foxcroft, C. (2013). Cross-cultural test adaptation, translation and tests in multiple languages. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed., pp. 102–125). Oxford University Press.

Dima, A. L. (2018). Scale validation in applied health research: Tutorial for a 6-step R-based psychometrics protocol. *Health Psychology and Behavioral Medicine*, *6*(1), 136–161. https://doi.org/10.1080/21642850.2018.1472602

Dorsey, S., Lucid, L., Martin, P., King, K. M., O'Donnell, K., Murray, L. K., Wasonga, A. I., Itemba, D. K., Cohen, J. A., Manongi, R., & Whetten, K. (2020). Effectiveness of task-shifted trauma-focused cognitive behavioral therapy for children who experienced parental death and posttraumatic stress in Kenya and Tanzania: A randomized clinical trial. *JAMA Psychiatry*, *77*(5), 464–473. https://doi.org/10.1001/jamapsychiatry.2019.4475

du Plessis, B., Kaminer, D., Hardy, A., & Benjamin, A. (2015). The contribution of different forms of violence exposure to internalizing and externalizing symptoms among young South African adolescents. *Child Abuse & Neglect*, *45*, 80–89. https://doi.org/10.1016/j.chiabu.2015.02.021

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Eggum, N. D., Sallquist, J., & Eisenberg, N. (2011). "Then it will be good": Negative life events and resilience in Ugandan youth. *Journal of Adolescent Research*, *26*(6), 766–796. https://doi.org/10.1177/0743558410391259

Familiar, I., Ruisenor-Escudero, H., Giordani, B., Bangirana, P., Nakasujja, N., Opoka, R., & Boivin, M. (2015). Use of the Behavior Rating Inventory of Executive Function and Child Behavior Checklist in Ugandan children with HIV or a history of severe malaria. *Journal of Developmental & Behavioral Pediatrics*, *36*(4), 277–284. https://doi.org/10.1097/DBP.0000000000000149

Fernald, L. C. H., Kariger, P., Engle, P. L., & Raikes, A. (2009). *Examining early child development in low-income countries: A toolkit for the assessment of children in the first five years of life*. The World Bank. https://elibrary.worldbank.org/doi/abs/10.1596/28107

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, *8*(4), 370–378. https://doi.org/10.1177/1948550617693063

Geibel, S., Habtamu, K., Mekonnen, G., Jani, N., Kay, L., Shibru, J., Bedilu, L., & Kalibala, S. (2016). Reliability and validity of an interviewer-administered adaptation of the Youth Self-Report for mental health screening of vulnerable young people in Ethiopia. *PLoS One*, *11*(2), Article e0147267. https://doi.org/10.1371/journal.pone.0147267

Gershoff, E. T., Grogan-Kaylor, A., Lansford, J. E., Chang, L., Zelli, A., Deater-Deckard, K., & Dodge, K. A. (2010). Parent discipline practices in an international sample: Associations with child behaviors and moderation by perceived normativeness. *Child Development*, *81*(2), 487–502. https://doi.org/10.1111/j.1467-8624.2009.01409.x

Gwandure, C. (2007). Home-based care for parents with AIDS: Impact on children's psychological functioning. *Journal of Child & Adolescent Mental Health*, *19*(1), 29–44. https://doi.org/10.2989/17280580709486633

Hall, B. J., Puffer, E., Murray, L. K., Ismael, A., Bass, J. K., Sim, A., & Bolton, P. A. (2014). The importance of establishing reliability and validity of assessment instruments for mental health problems: An example from Somali children and adolescents living in three refugee camps in Ethiopia. *Psychological Injury and Law*, *7*(2), 153–164. https://doi.org/10.1007/s12207-014-9188-9

Harder, V. S., Mutiso, V. N., Khasakhala, L. I., Burke, H. M., Rettew, D. C., Ivanova, M. Y., & Ndetei, D. M. (2014). Emotional and behavioral problems among impoverished Kenyan youth: Factor structure and sex-differences. *Journal of Psychopathology and Behavioral Assessment*, *36*(4), 580–590. https://doi.org/10.1007/s10862-014-9419-0

Hoosen, N., Davids, E. L., de Vries, P. J., & Shung-King, M. (2018). The Strengths and Difficulties Questionnaire (SDQ) in Africa: A scoping review of its application and validation. *Child and Adolescent Psychiatry and Mental Health*, *12*(1), Article 6. https://doi.org/10.1186/s13034-017-0212-1

Isaksson, J., Deyessa, N., Berhane, Y., & Högberg, U. (2017). Early adversity and psychiatric symptoms – a prospective study on Ethiopian mothers and their children. *BMC Psychiatry*, *17*, Article 344. https://doi.org/10.1186/s12888-017-1500-2

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Dumenci, L., Almqvist, F., Bilenberg, N., Bird, H., Broberg, A. G., Dobrean, A., Döpfner, M., Erol, N., Forns, M., Hannesdottir, H., Kanbayashi, Y., Lambert, M. C., Leung, P., Minaei, A., Mulatu, M. S., Novik, T., … Verhulst, F. C. (2007). The generalizability of the Youth Self-Report syndrome structure in 23 societies. *Journal of Consulting and Clinical Psychology*, *75*(5), 729–738. https://doi.org/10.1037/0022-006X.75.5.729

Ivanova, M. Y., Dobrean, A., Dopfner, M., Erol, N., Fombonne, E., Fonseca, A. C., Frigerio, A., Grietens, H., Hannesdottir, H., Kanbayashi, Y., Lambert, M., Achenbach, T. M., Larsson, B., Leung, P., Liu, X., Minaei, A., Mulatu, M. S., Novik, T. S., Oh, K. J., … Chen, W. J. (2007). Testing the 8-syndrome structure of the Child Behavior Checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology*, *36*(3), 405–417. https://doi.org/10.1080/15374410701444363

Jenkins, R., Baingana, F., Belkin, G., Borowitz, M., Daly, A., Francis, P., Friedman, J., Garrison, P., Kauye, F., Kiima, D., Mayeya, J., Mbatia, J., Tyson, S., Njenga, F., Gureje, O., & Sadiq, S. (2010). Mental health and the development agenda in sub-Saharan Africa. *Psychiatric Services*, *61*(3), 229–234. https://doi.org/10.1176/ps.2010.61.3.229

Kariuki, S. M., Abubakar, A., Murray, E., Stein, A., & Newton, C. R. J. C. (2016). Evaluation of psychometric properties and factorial structure of the pre-school child behaviour checklist at the Kenyan Coast. *Child and Adolescent Psychiatry and Mental Health*, *10*, Article 1. https://doi.org/10.1186/s13034-015-0089-9

Klasen, F., Oettingen, G., Daniels, J., & Adam, H. (2010). Multiple trauma and mental health in former Ugandan child soldiers. *Journal of Traumatic Stress*, *23*(5), 573–581. https://doi.org/10.1002/jts.20557

Kumar, M., Amugune, B., Madeghe, B., Wambua, G. N., Osok, J., Polkonikova-Wamoto, A., Bukusi, D., Were, F., & Huang, K.-Y. (2018). Mechanisms associated with maternal adverse childhood experiences on offspring's mental health in Nairobi informal settlements: A mediational model testing approach. *BMC Psychiatry*, *18*, Article 381. https://doi.org/10.1186/s12888-018-1953-y

Kusi-Mensah, K., Donnir, G., Wemakor, S., Owusu-Antwi, R., & Omigbodun, O. (2019). Prevalence and patterns of mental disorders among primary school age children in Ghana: Correlates with academic achievement. *Journal of Child & Adolescent Mental Health*, *31*(3), 214–223. https://doi.org/10.2989/17280583.2019.1678477

LeCroix, R. H., Chan, W. Y., Henrich, C., Palin, F., Shanley, J., & Armistead, L. (2020). Maternal HIV and adolescent functioning in South Africa: The role of the mother-child relationship. *The Journal of Early Adolescence*, *40*(1), 83–103. https://doi.org/10.1177/0272431618824726

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Magai, D. N., Koot, H. M., Newton, C. R., & Abubakar, A. (2021). Long-term mental health and quality of life outcomes of neonatal insults in Kilifi, Kenya. *Child Psychiatry & Human Development*. Advance online publication. https://doi.org/10.1007/s10578-020-01079-1

Magai, D. N., Malik, J. A., & Koot, H. M. (2018). Emotional and behavioral problems in children and adolescents in Central Kenya. *Child Psychiatry & Human Development*, *49*(4), 659–671. https://doi.org/10.1007/s10578-018-0783-y

Malcolm-Smith, S., Woolley, D., & Ward, C. L. (2015). Examining empathy and its association with aggression in young Western Cape children. *Journal of Child and Adolescent Mental Health*, *27*(2), 135–147. https://doi.org/10.2989/17280583.2015.1085386

Meinck, F., Boyes, M. E., Cluver, L., Ward, C. L., Schmidt, P., DeStone, S., & Dunne, M. P. (2018). Adaptation and psychometric properties of the ISPCAN Child Abuse Screening Tool for use in trials (ICAST-Trial) among South African adolescents and their primary caregivers. *Child Abuse & Neglect*, *82*, 45–58. https://doi.org/https://doi.org/10.1016/j.chiabu.2018.05.022

Meinck, F., Cluver, L. D., Orkin, F. M., Kuo, C., Sharma, A. D., Hensels, I. S., & Sherr, L. (2017). Pathways from family disadvantage via abusive parenting and caregiver mental health to adolescent health risks in South Africa. *Journal of Adolescent Health*, *60*(1), 57–64. https://doi.org/10.1016/j.jadohealth.2016.08.016

Meinck, F., Orkin, F. M., & Cluver, L. (2019). Does free schooling affect pathways from adverse childhood experiences via mental health distress to HIV risk among adolescent girls in South Africa: A longitudinal moderated pathway model. *Journal of the International AIDS Society*, *22*, Article e25262. https://doi.org/https://doi.org/10.1002/jia2.25262

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., 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. https://doi.org/10.1007/s11136-017-1765-4

Mokkink, L. B., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & 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

Mokkink, L. B., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2019). *COSMIN study design checklist for patient-reported outcome measurement instruments*. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf

Mulatu, M. S. (1995). Prevalence and risk factors of psychopathology in Ethiopian children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*(1), 100–109. https://doi.org/10.1097/00004583-199501000-00020

Mulatu, M. S. (1997). *Stress, coping, and adaptation in Ethiopian adolescents: Testing the applicability of the transactional model of stress and coping*. [Unpublished doctoral thesis]. Queen's University.

Murray, L. K., Hall, B. J., Dorsey, S., Ugueto, A. M., Puffer, E. S., Sim, A., Ismael, A., Bass, J., Akiba, C., Lucid, L., Harrison, J., Erikson, A., & Bolton, P. A. (2018). An evaluation of a common elements treatment approach for youth in Somali refugee camps. *Global Mental Health*, *5*, Article e16. https://doi.org/10.1017/gmh.2018.7

Murray, S. M., Bolton, P., Kane, J. C., Lakin, D. P., Skavenski Van Wyk, S., Paul, R., & Murray, L. K. (2020). Measuring symptoms of psychopathology in Zambian orphans and vulnerable children: Scale validation and psychometric evaluation. *Assessment*, *27*(6), 1335–1348. https://doi.org/10.1177/1073191118780455

Ndetei, D., Mutiso, V., Musyimi, C., Mokaya, A., Anderson, K., McKenzie, K., Musau, A., Ndetei, D. M., Mokaya, A. G., & Anderson, K. K. (2016). The prevalence of mental disorders among upper primary school children in Kenya. *Social Psychiatry & Psychiatric Epidemiology*, *51*(1), 63–71. https://doi.org/10.1007/s00127-015-1132-0

Nezafat Maldonado, B., Chandna, J., & Gladstone, M. (2019). A systematic review of tools used to screen and assess for externalising behaviour symptoms in low and middle income settings. *Global Mental Health*, *6*, Article e13. https://doi.org/10.1017/gmh.2019.11

Ng, L. C., Kirk, C. M., Kanyanganzi, F., Fawzi, M. C. S., Sezibera, V., Shema, E., Bizimana, J. I., Cyamatare, F. R., & Betancourt, T. S. (2015). Risk and protective factors for suicidal ideation and behaviour in Rwandan children. *British Journal of Psychiatry*, *207*(3), 262–268. https://doi.org/10.1192/bjp.bp.114.154591

Nsabimana, E., Rutembesa, E., Wilhelm, P., & Martin-Soelch, C. (2019). Effects of institutionalization and parental living status on children's self-esteem, and externalizing and internalizing problems in Rwanda. *Frontiers in Psychiatry*, *10*, Article 442. https://doi.org/10.3389/fpsyt.2019.00442

Osman, F., Flacking, R., Schön, U.-K., & Klingberg-Allvin, M. (2017). A support program for Somali-born parents on children's behavioral problems. *Pediatrics*, *139*(3), Article e20162764. https://doi.org/10.1542/peds.2016-2764

Palin, F. L., Armistead, L., Clayton, A., Ketchen, B., Lindner, G., Kokot-Louw, P., & Pauw, A. (2009). Disclosure of maternal HIV-infection in South Africa: Description and relationship to child functioning. *AIDS and Behavior*, *13*(6), 1241–1252. https://doi.org/10.1007/s10461-008-9447-4

Peltzer, K., & Pengpid, S. (2013). Intimate partner violence, mental problems and behavioural problems among pre-school children in Vhembe district, South Africa. *Child Abuse Research in South Africa*, *14*(1), 14–21.

Pienaar, E., Grobler, L., Busgeeth, K., Eisinga, A., & Siegfried, N. (2011). Developing a geographic search filter to identify randomised controlled trials in Africa: Finding the optimal balance between sensitivity and precision. *Health Information & Libraries Journal*, *28*(3), 210–215. https://doi.org/10.1111/j.1471-1842.2011.00936.x

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. https://doi.org/10.1007/s11136-018-1798-3

Rasmussen, A., Cissé, A., Han, Y., & Roubeni, S. (2018). Migration factors in West African immigrant parents' perceptions of their children's neighborhood safety. *American Journal of Community Psychology*, *61*(3–4), 321–331. https://doi.org/10.1002/ajcp.12230

Rawatlal, N., Kliewer, W., & Pillay, B. J. (2015). Adolescent attachment, family functioning and depressive symptoms. *South African Journal of Psychiatry*, *21*(3), 80–85. https://doi.org/10.4102/sajpsychiatry.v21i3.672

Rochat, T. J., Arteche, A. X., Stein, A., Mitchell, J., & Bland, R. M. (2015). Maternal and child psychological outcomes of HIV disclosure to young children in rural South Africa: The Amagugu intervention. *AIDS*, *29*, S67–S79. https://doi.org/10.1097/QAD.0000000000000668

Rochat, T. J., Houle, B., Stein, A., Coovadia, H., Coutsoudis, A., Desmond, C., Newell, M.-L., & Bland, R. M. (2016). Exclusive breastfeeding and cognition, executive function, and behavioural disorders in primary school-aged children in rural South Africa: A cohort analysis. *PLOS Medicine*, *13*(6), Article e1002044. https://doi.org/10.1371/journal.pmed.1002044

Rochat, T. J., Houle, B., Stein, A., Pearson, R. M., Newell, M. L., & Bland, R. M. (2017). Cohort profile: The Siyakhula cohort, rural South Africa. *International Journal of Epidemiology*, *46*(6), 1755-1756n. https://doi.org/10.1093/ije/dyx148

Rochat, T. J., Mitchell, J., Lubbe, A. M., Stein, A., Tomlinson, M., & Bland, R. M. (2017). Communication about HIV and death: Maternal reports of primary school-aged children's questions after maternal HIV disclosure in rural South Africa. *Social Science & Medicine*, *172*, 124–134. https://doi.org/10.1016/j.socscimed.2016.10.031

Ruiseñor-Escudero, H., Familiar, I., Nakasujja, N., Bangirana, P., Opoka, R., Giordani, B., & Boivin, M. (2015). Immunological correlates of behavioral problems in school-aged children living with HIV in Kayunga, Uganda. *Global Mental Health*, *2*, Article e9. https://doi.org/10.1017/gmh.2015.7

Sharp, C., Venta, A., Marais, L., Skinner, D., Lenka, M., & Serekoane, J. (2014). First evaluation of a population-based screen to detect emotional-behavior disorders in orphaned children in sub-Saharan Africa. *AIDS and Behavior*, *18*(6), 1174–1185. https://doi.org/10.1007/s10461-014-0739-6

Shenderovich, Y., Cluver, L., Eisner, M., & Murray, A. L. (2020). Moderators of treatment effects in a child maltreatment prevention programme in South Africa. *Child Abuse & Neglect*, *106*, Article 104519. https://doi.org/https://doi.org/10.1016/j.chiabu.2020.104519

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/S11336-008-9101-0

Sipsma, H., Eloff, I., Makin, J., Finestone, M., Ebersohn, L., Visser, M. J., Sikkema, K. J., Allen, C. A. B., Ferreira, R., & Forsyth, B. (2013). Behavior and psychological functioning of young children of HIV-positive mothers in South Africa. *AIDS Care*, *25*(6), 721–725. https://doi.org/10.1080/09540121.2013.779627

Skinner, A. T., Oburu, P., Lansford, J. E., & Bacchini, D. (2014). Childrearing violence and child adjustment after exposure to Kenyan post-election violence. *Psychology of Violence*, *4*(1), 37–50. https://doi.org/10.1037/a0033237

Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, *17*(2), 268–274. https://doi.org/10.1111/j.1365-2753.2010.01434.x

Swain, K. D., Pillay, B. J., & Kliewer, W. (2017). Traumatic stress and psychological functioning in a South African adolescent community sample. *South African Journal of Psychiatry*, *23*, Article a1008. https://doi.org/10.4102/sajpsychiatry.v23i0.1008

Sweetland, A. C., Belkin, G. S., & Verdeli, H. (2014). Measuring depression and anxiety in sub-Saharan Africa. *Depression and Anxiety*, *31*(3), 223–232. https://doi.org/10.1002/da.22142

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., de Vet, H. C. W., Bouter, L. M., Alonso, J., Westerman, M. J., Patrick, D. L., & Mokkink, L. B. (2018). *COSMIN methodology for assessing the content validity of PROMs: User manual*. https://cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf

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. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research*, *27*(5), 1159–1170. https://doi.org/10.1007/s11136-018-1829-0

Thornton, V. J., Asanbe, C. B., & Denton, E. D. (2019). Clinical risk factors among youth at high risk for suicide in South Africa and Guyana. *Depression and Anxiety*, *36*(5), 423–432. https://doi.org/10.1002/da.22889

United Nations Statistics Division. (1998). *United Nations Standard Country Code (M49)*. United Nations. https://unstats.un.org/unsd/methodology/m49/

Vacha-Haase, T., & Thompson, B. (2011). Score reliability: A retrospective look back at 12 years of reliability generalization studies. *Measurement and Evaluation in Counseling and Development*, *44*(3), 159–168. https://doi.org/10.1177/0748175611409845

Van De Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17–45). Cambridge University Press.

Van Eeden, R., & Mantsha, T. R. (2007). Theoretical and methodological considerations in the translation of the 16PF5 into an African language. *South African Journal of Psychology*, *37*(1), 62–81. https://doi.org/10.1177/008124630703700105

van Westrhenen, N., Fritz, E., Vermeer, A., Boelen, P., & Kleber, R. (2019). Creative arts in psychotherapy for traumatized children in South Africa: An evaluation study. *PLoS One*, *14*(2), Article e0210857. https://doi.org/10.1371/journal.pone.0210857

Van Widenfelt, B. M., Treffers, P. D. A., De Beurs, E., Siebelink, B., Koudijs, E., & Siebelink, B. M. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child & Family Psychology Review*, *8*(2), 135–147. https://doi.org/10.1007/s10567-005-4752-1

Visser, M. J., Hecker, H. E., & Jordaan, J. (2018). A comparative study of the psychological problems of HIV-infected and HIV-uninfected children in a South African sample. *AIDS Care*, *30*(5), 596–603. https://doi.org/10.1080/09540121.2017.1417530

Wadji, D. L., Ketcha Wanda, G. J. M., Wicky, C., Morina, N., & Martin-Soelch, C. (2020). From the mother to the child: The intergenerational transmission of experiences of violence in mother-child dyads exposed to intimate partner violence in Cameroon. *Journal of Interpersonal Violence*. Advance online publication. https://doi.org/10.1177/0886260520948148

Zumbo, B. D., & Chan, E. K. H. (2014). Setting the stage for validity and validation in social, behavioral, and health sciences: Trends in validation practices. In B. D. Zumbo & E. K. H. Chan (Eds.), *Validity and validation in social, behavioral, and health sciences* (pp. 3–8). Springer International Publishing.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. https://doi.org/10.22237/jmasm/1177992180

Supplementary Materials

Table S1: *Summary of Databases, Search Terms, and Number of Studies Pulled from Each Database*

Table S2: *Excluded Studies and Reasons for Exclusion (n = 64)*

Appendix 1: *Journal articles with ASEBA Forms, a Sub-Saharan African sample, but no Psychometric Properties (n = 87)*

Table S1
Summary of Databases, Search Terms, and Number of Studies Pulled from Each Database

Database	Search terms	Notes	Number of studies	
PubMed	(((“child behaviour checklist” OR “child behavior checklist” OR CBCL OR “youth self-report” OR “teacher’s report form”)) AND (Africa OR African OR Angola OR Benin OR Botswana OR “Burkina Faso” OR Burundi OR Cameroon OR “Central African Republic” OR Chad OR Congo OR “Democratic Republic of Congo” OR “Republic of Congo” OR Djibouti OR “Equatorial Guinea” OR Eritrea OR Eswatini OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR “Guinea Bissau” OR “Ivory Coast” OR “Cote d’Ivoire” OR Kenya OR Lesotho OR Liberia OR Malawi OR Mali OR Mauritania OR Mozambique OR Mocambique OR Namibia OR Niger OR Nigeria OR Rwanda OR Senegal OR “Sierra Leone” OR Somalia OR “South Africa” OR Sudan OR South Sudan OR Swaziland OR Tanzania OR Togo OR Uganda OR Zaire OR Zambia OR Zimbabwe OR “sub-Saharan Africa” OR “sub-Saharan African”)) NOT (“African American” OR “African-American”)	“All fields” selected for all three search boxes. No filters or restrictions were applied (e.g., date, document type, full text availability).		129
EBSCOhost Academic Search Premier Africa-Wide Information Health Source: Nursing/Academic Edition CINAHL ERIC APA PsycInfo APA PsycArticles	TX (“child behaviour checklist” OR “child behavior checklist” OR CBCL) AND TX (achenbach or ASEBA) AND TX (Africa OR Angola OR Benin OR Botswana OR “Burkina Faso” OR Burundi OR Cameroon OR “Central African Republic” OR Chad OR Congo OR “Democratic Republic of Congo” OR “Republic of Congo” OR Djibouti OR “Equatorial Guinea” OR Eritrea OR Eswatini OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR “Guinea Bissau” OR “Ivory Coast” OR “Cote d’Ivoire” OR Kenya OR Lesotho OR Liberia OR Malawi OR Mali OR Mauritania OR Mozambique OR Mocambique OR Namibia OR Niger OR Nigeria OR Rwanda OR Senegal OR “Sierra Leone” OR Somalia OR “South Africa” OR Sudan OR South Sudan OR Swaziland OR Tanzania OR	“All text” fields were selected for all search boxes. The initial search yielded too many results (1745). We narrowed the search by adding another line to the search (achenbach OR ASEBA). No filters or restrictions were applied.	Total Academic Search Africa-Wide Health Source CINAHL ERIC APA PsycInfo APA PsycArticles	853 512 17 96 45 1 62 120

	Togo OR Uganda OR Zaire OR Zambia OR Zimbabwe OR “sub-Saharan Africa” OR “sub-Saharan African”) NOT TX (“African American” OR “African-American”)		
Scopus	ALL ("Child behaviour checklist" OR "Child behavior checklist" OR cbcl OR "youth self-report" OR "teacher's report form") AND ALL (achenbach OR aseba) AND ALL (Africa OR African OR "sub-Saharan Africa" OR "sub-Saharan AND African" OR Angola OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR Cameroon OR "Central African Republic" OR chad OR Congo OR "Democratic Republic of Congo" OR "Republic of Congo" OR Djibouti OR "Equatorial Guinea" OR Eritrea OR Eswatini OR Ethiopia OR Gabon OR Gambia OR Ghana OR guinea OR "Guinea Bissau" OR "Ivory Coast" OR "Cote d'Ivoire" OR Kenya OR Lesotho OR Liberia OR Malawi OR Mali OR Mauritania OR Mozambique OR Mocambique OR Namibia OR Niger OR Nigeria OR Rwanda OR Senegal OR "Sierra Leone" OR Somalia OR "South Africa" OR Sudan OR "South Sudan" OR Swaziland OR Tanzania OR Togo OR Uganda OR Zaire OR Zambia OR Zimbabwe)) AND NOT ALL ("African American" OR "African-American") AND NOT INDEX (medline)	Results from MEDLINE were filtered out to avoid duplicates of the PubMed search results. The initial search yielded too many results (1288) with the phrase “youth self-report” that were not relevant to the ASEBA forms. We therefore narrowed the search by adding (achenbach OR ASEBA). No other filters or restrictions were applied.	204
Google Scholar	("child behaviour checklist" OR "child behaviour checklist" OR cbcl OR “youth self-report” OR “teacher’s report form”) AND (Africa OR African OR "sub-Saharan Africa" OR “sub-Saharan African”) -"African American"	We included all results from the first twenty pages (20 x 10 results per page), as the results became less relevant after this point. We simplified the SSA search terms due to space limits in the search box.	200
Proquest	("child behaviour checklist" OR "child behavior checklist" OR “youth self-report” or “teacher’s report form”) AND	The initial search looked for search terms “Anywhere”, but this yielded too many results (1261). Restricting the SSA	37

	ab(Africa OR African OR Angola OR Benin OR Botswana OR "Burkina Faso" OR Burundi OR Cameroon OR "Central African Republic" OR Chad OR Congo OR "Democratic Republic of Congo" OR "Republic of Congo" OR Djibouti OR "Equatorial Guinea" OR Eritrea OR Eswatini OR Ethiopia OR Gabon OR Gambia OR Ghana OR Guinea OR "Guinea Bissau" OR "Ivory Coast" OR "Cote d'Ivoire" OR Kenya OR Lesotho OR Liberia OR Malawi OR Mali OR Mauritania OR Mozambique OR Mocambique OR Namibia OR Niger OR Nigeria OR Rwanda OR Senegal OR "Sierra Leone" OR Somalia OR "South Africa" OR Sudan OR "South Sudan" OR Swaziland OR Tanzania OR Togo OR Uganda OR Zaire OR Zambia OR Zimbabwe OR "sub-Saharan Africa" OR "sub-Saharan African") NOT ("African American" OR "African-American")	search terms to the 'Location' search field yielded too few results (4). The final search restricted the SSA search terms to "Abstract". All search results were limited to books, book chapters, conference papers and proceedings, and thesis/dissertations. Only 1 result was a book and the rest were theses.	
PRIMO University of Cape Town Libraries Catalogue	('Child behaviour checklist' OR 'child behavior checklist') AND (africa OR african OR 'sub-Saharan Africa' OR 'sub-saharan African') NOT ('african american' OR 'african-american')	Results were filtered to include books, book chapters, dissertations and conference proceedings. The full search terms used for other databases yielded too many results (4795). We modified the SSA search terms but there were still too many results (3271). Finally, we simplified the ASEBA terms to "child behaviour checklist" OR "child behavior checklist" only.	169

Note. CBCL = Child Behaviour Checklist, ASEBA = Achenbach System of Empirically Based Assessment, SSA = sub-Saharan Africa.

Table S2

Excluded Studies and Reasons for Exclusion (n = 64)

Reference	Country	Reason for exclusion
East Africa (*n* = 20)		
Aramburu Alegret, I., Pérez-Testor, C., Mercadal Rotger, J., Salamero Baró, M., Davins Pujols, M., Mirabent Junyent, V., Aznar Martínez, B., & Brodzinsky, D. (2020). Influence of communicative openness on the psychological adjustment of internationally adopted adolescents. *Journal of Research on Adolescence*, *30*, 226–237. https://doi.org/10.1111/jora.12464	Ethiopia (Spain)	No stratified data for Ethiopian participants
Bjørknes, R., Manger, T., Bjørknes, R., & Manger, T. (2013). Can parent training alter parent practice and reduce conduct problems in ethnic minority children? A randomized controlled trial. *Prevention Science*, *14*(1), 52–63. https://doi.org/10.1007/s11121-012-0299-9	Somalia (Norway)	No stratified data for Somali participants
Bornstein, M. H., Putnick, D. L., Lansford, J. E., Al-Hassan, S. M., Bacchini, D., Bombi, A. S., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Tirado, L. M. U., Zelli, A., & Alampay, L. P. (2017). "Mixed blessings": Parental religiousness, parenting, and child adjustment in global perspective. *Journal of Child Psychology & Psychiatry*, *58*(8), 880–892. https://doi.org/10.1111/jcpp.12705	Kenya	No stratified data for Kenyan participants
Cederblad, M., & Hook, B. (1999). Mental health in international adoptees as teenagers and young adults: An epidemiological study. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *40*(8), 1239. https://doi.org/10.1111/1469-7610.00540	Ethiopia (Sweden)	No stratified data for Ethiopian participants
Chang, L., Lu, H. J., Lansford, J. E., Skinner, A. T., Bornstein, M. H., Steinberg, L., Dodge, K. A., Chen, B. Bin, Tian, Q., Bacchini, D., Deater-Deckard, K., Pastorelli, C., Alampay, L. P., Sorbring, E., Al-Hassan, S. M., Oburu, P., Malone, P. S., Di Giunta, L., Tirado, L. M. U., & Tapanya, S. (2019). Environmental harshness and unpredictability, life history, and social and academic behavior of adolescents in nine countries. *Developmental Psychology*, *55*(4), 890–903. https://doi.org/10.1037/dev0000655	Kenya	No stratified data for Kenyan participants
Chaudhury, S., Brown, F. L., Kirk, C. M., Mukunzi, S., Nyirandagijimana, B., Mukandanga, J., Ukundineza, C., Godfrey, K., Ng, L. C., Brennan, R. T., & Betancourt, T. S. (2016). Exploring the potential of a family-based prevention intervention to reduce alcohol use and violence within HIV-affected families in Rwanda. *AIDS Care*, *28*, 118–129. https://doi.org/10.1080/09540121.2016.1176686	Rwanda	Non-standard use
Deater-Deckard, K., Godwin, J., Lansford, J. E., Bacchini, D., Bombi, A. S., Bornstein, M. H., Chang, L., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Alampay, L. P., Uribe Tirado, L. M., Zelli, A., & Al-Hassan, S. M. (2018). Within- and between-person and group variance in behavior and beliefs in cross-cultural longitudinal data. *Journal of Adolescence*, *62*, 207–217. https://doi.org/10.1016/j.adolescence.2017.06.002	Kenya	No stratified data for Kenyan participants

Eggum-Wilkens, N. D., Zhang, L., & Farago, F. (2017). Karamojong adolescents in Tororo, Uganda: Life events, adjustment problems, and protective factors. *Journal of Loss & Trauma*, *22*(4), 283–296. https://doi.org/10.1080/15325024.2017.1284508	Uganda	Non-standard use
Elovainio, M., Hakulinen, C., Pulkki-Råback, L., Raaska, H., & Lapinleimu, H. (2018). The network structure of childhood psychopathology in international adoptees. *Journal of Child & Family Studies*, *27*(7), 2161–2170. https://doi.org/10.1007/s10826-018-1046-z	Ethiopia and Mozambique (Finland)	No stratified data for Ethiopian or Mozambican participants
Elovainio, M., Raaska, H., Sinkkonen, J., Mäkipää, S., & Lapinleimu, H. (2015). Associations between attachment-related symptoms and later psychological problems among international adoptees: Results from the FinAdo study. *Scandinavian Journal of Psychology*, *56*(1), 53–61. https://doi.org/10.1111/sjop.12174	Ethiopia and Mozambique (Finland)	No stratified data for Ethiopian or Mozambican participants
Kane, J. C., Murray, L. K., Bass, J. K., Johnson, R. M., & Bolton, P. (2016). Validation of a substance and alcohol use assessment instrument among orphans and vulnerable children in Zambia using Audio Computer Assisted Self-Interviewing (ACASI). *Drug and Alcohol Dependence*, *166*, 85–92. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2016.06.026	Zambia	Non-standard use
Lansford, J. E., Chang, L., Dodge, K. A., Malone, P. S., Oburu, P., Palmérus, K., Bacchini, D., Pastorelli, C., Bombi, A. S., Zelli, A., Tapanya, S., Chaudhary, N., Deater-Deckard, K., Manke, B., & Quinn, N. (2005). Physical discipline and children's adjustment: Cultural normativeness as a moderator. *Child Development*, *76*(6), 1234–1246. https://doi.org/10.1111/j.1467-8624.2005.00847.x	Kenya	No stratified ASEBA data for Kenyan participants
Lim, S.-L., & Ogawa, Y. (2014). "Once I had kids, now I am raising kids": Child-Parent Relationship Therapy (CPRT) with a Sudanese refugee family—A case study. *International Journal of Play Therapy*, *23*(2), 70–89. https://doi.org/10.1037/a0036362	Sudan	Case study
Mulatu, M. S. (1995). Prevalence and risk factors of psychopathology in Ethiopian children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*(1), 100–109. https://doi.org/10.1097/00004583-199501000-00020	Ethiopia	Non-standard use
Nduwimana, E., Mukunzi, S., Ng, L. C., Kirk, C. M., Bizimana, J. I., & Betancourt, T. S. (2017). Mental health of children living in foster families in rural Rwanda: The role of HIV and the family environment. *AIDS and Behavior*, *21*(6), 1518–1529. https://doi.org/10.1007/s10461-016-1482-y	Rwanda	Non-standard use
Putnick, D. L., Bornstein, M. H., Lansford, J. E., Malone, P. S., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Tirado, L. M. U., Zelli, A., Alampay, L. P., Al-Hassan, S. M., Bacchini, D., Bombi, A. S., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., & Oburu, P. (2015). Perceived mother and father acceptance-rejection predict four unique aspects of child adjustment across nine countries. *Journal of Child Psychology and Psychiatry*, *56*(8), 923–932. https://doi.org/10.1111/jcpp.12366	Kenya	No stratified data for Kenyan participants
Rita, N., Elovainio, M., Raaska, H., Lahti-Nuuttila, P., Matomäki, J., Sinkkonen, J., & Lapinleimu, H. (2017). Child and family-related predictors of psychological outcomes in children adopted from abroad; what is the role of caregiver time? *Scandinavian Journal of Psychology*, *58*(4), 312–317. https://doi.org/10.1111/sjop.12374	Ethiopia and Mozambique (Finland)	No stratified data for Ethiopian or Mozambican participants

Roskam, I., Stievenart, M., Tessier, R., Muntean, A., Escobar, M., Santelices, M., Juffer, F., Ijzendoorn, M., & Pierrehumbert, B. (2014). Another way of thinking about ADHD: The predictive role of early attachment deprivation in adolescents' level of symptoms. *Social Psychiatry & Psychiatric Epidemiology*, *49*(1), 133–144. https://doi.org/10.1007/s00127-013-0685-z	Ethiopia	No stratified data for Ethiopian participants
Schenck-Fontaine, A., Lansford, J. E., Skinner, A. T., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Oburu, P., Pastorelli, C., Sorbring, E., & Steinberg, L. (2020). Associations between perceived material deprivation, parents' discipline practices, and children's behavior problems: An international perspective. *Child Development*, *91*(1), 307–326. https://doi.org/https://doi.org/10.1111/cdev.13151	Kenya	No stratified data for Kenyan participants
Vollebergh, W. A. M., Have, M., Dekovic, M., Oosterwegel, A., Pels, T., Veenstra, R., Winter, A., Ormel, H., Verhulst, F., ten Have, M., Veenstra, R., & de Winter, A. (2005). Mental health in immigrant children in the Netherlands. *Social Psychiatry & Psychiatric Epidemiology*, *40*(6), 489–496. https://doi.org/10.1007/s00127-005-0906-1	Somalia (The Netherlands)	No stratified data for Somali participants
Southern Africa (*n* = 19)		
Barbarin, O. A., & Richter, L. (2001). Economic status, community danger and psychological problems among South African children. *Childhood*, *8*(1), 115–133. https://doi.org/10.1177/0907568201008001007	South Africa	Non-standard use
Barone, L., Lionetti, F., & Green, J. (2017). A matter of attachment? How adoptive parents foster post-institutionalized children's social and emotional adjustment. *Attachment & Human Development*, *19*(4), 323–339. https://doi.org/10.1080/14616734.2017.1306714	South Africa	No stratified data for South Africa participants
Cluver, L. D., Orkin, F. M., Campeau, L., Toska, E., Webb, D., Carlqvist, A., & Sherr, L. (2019). Improving lives by accelerating progress towards the UN Sustainable Development Goals for adolescents living with HIV: a prospective cohort study. *The Lancet Child & Adolescent Health*, *3*(4), 245–254. https://doi.org/10.1016/s2352-4642(19)30033-1	South Africa	Non-standard use
Cluver, L. D., Rudgard, W. E., Toska, E., Zhou, S., Campeau, L., Shenderovich, Y., Orkin, M., Desmond, C., Butchart, A., & Taylor, H. (2020). Violence prevention accelerators for children and adolescents in South Africa: A path analysis using two pooled cohorts. *PLoS Medicine*, *17*(11), e1003383. https://doi.org/10.1371/journal.pmed.1003383	South Africa	Non-standard use
Davies, L. C., & McKelvey, R. S. (1998). Emotional and behavioural problems and competencies among immigrant and non-immigrant adolescents. *Australian & New Zealand Journal of Psychiatry*, *32*(5), 658–665. https://doi.org/10.3109/00048679809113120	South Africa	No stratified data for South Africa participants
de la Loge, C., Hunter, S. J., Schiemann, J., & Yang, H. (2010). Assessment of behavioral and emotional functioning using standardized instruments in children and adolescents with partial-onset seizures treated with adjunctive levetiracetam in a randomized, placebo-controlled trial. *Epilepsy & Behavior*, *18*(3), 291–298. https://doi.org/10.1016/j.yebeh.2010.04.017	South Africa	No stratified data for South Africa participants

Diliberto, R., & Kearney, C. A. (2018). Latent class symptom profiles of selective mutism: Identification and linkage to temperamental and social constructs. *Child Psychiatry and Human Development*, *49*(4), 551–562. https://doi.org/10.1007/s10578-017-0774-4	South Africa	No stratified data for South Africa participants
Gardner, F., Waller, R., Maughan, B., Cluver, L., & Boyes, M. (2015). What are the risk factors for antisocial behavior among low-income youth in Cape Town? *Social Development*, *24*(4), 798–814. https://doi.org/10.1111/sode.12127	South Africa	Non-standard use
Gautam, P., Lebel, C., Narr, K. L., Mattson, S. N., May, P. A., Adnams, C. M., Riley, E. P., Jones, K. L., Kan, E. C., & Sowell, E. R. (2015). Volume changes and brain-behavior relationships in white matter and subcortical gray matter in children with prenatal alcohol exposure. *Human Brain Mapping*, *36*(6), 2318–2329. https://doi.org/10.1002/hbm.22772	South Africa	No stratified data for South Africa participants
Govender, K., Reardon, C., Quinlan, T., & George, G. (2014). Children's psychosocial wellbeing in the context of HIV/AIDS and poverty: a comparative investigation of orphaned and non-orphaned children living in South Africa. *BMC Public Health*, *14*(1), 998–1021. https://doi.org/10.1186/1471-2458-14-615	South Africa	Non-standard use
Hoare, J., Stein, D. J., Heany, S. J., Fouche, J.-P., Phillips, N., Er, S., Myer, L., Zar, H. J., Horvath, S., & Levine, A. J. (2020). Accelerated epigenetic aging in adolescents from low-income households is associated with altered development of brain structures. *Metabolic Brain Disease*, *35*(8), 1287–1298. https://doi.org/10.1007/s11011-020-00589-0	South Africa	Non-standard use
Nkomo, P., Mathee, A., Naicker, N., Galpin, J., Richter, L. M., & Norris, S. A. (2017). The association between elevated blood lead levels and violent behavior during late adolescence: The South African Birth to Twenty Plus cohort. *Environment International*, *109*, 136–145. https://doi.org/https://doi.org/10.1016/j.envint.2017.09.004	South Africa	ASEBA data not reported
Nkomo, P., Naicker, N., Mathee, A., Galpin, J., Richter, L. M., & Norris, S. A. (2018). The association between environmental lead exposure with aggressive behavior, and dimensionality of direct and indirect aggression during mid-adolescence: Birth to Twenty Plus cohort. *Science of the Total Environment*, *612*, 472–479. https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.08.138	South Africa	Non-standard use
Nyamukapa, C. A., Gregson, S., Lopman, B., Saito, S., Watts, H. J., Monasch, R., & Jukes, M. C. H. (2008). HIV-Associated orphanhood and children's psychosocial distress: Theoretical framework tested with data from Zimbabwe. *American Journal of Public Health*, *98*(1), 133–141. https://doi.org/10.2105/AJPH.2007.116038	Zimbabwe	Non-standard use
Reardon, C., George, G., Mucheuki, C., Govender, K., & Quinlan, T. (2015). Psychosocial and health risk outcomes among orphans and non-orphans in mixed households in KwaZulu-Natal, South Africa. *African Journal of AIDS Research*, *14*(4), 323–331. https://doi.org/10.2989/16085906.2015.1095774	South Africa	Non-standard use
Spence, S. H., Donovan, C. L., March, S., Gamble, A., Anderson, R. E., Prosser, S., & Kenardy, J. (2011). A randomized controlled trial of online versus clinic-based CBT for adolescent anxiety. *Journal of Consulting and Clinical Psychology*, *79*(5), 629–642. https://doi.org/10.1037/a0024512	South Africa (Australia)	No stratified data for South Africa participants

Toska, E., Cluver, L., Orkin, M., Bains, A., Sherr, L., Berezin, M., & Gulaid, L. (2019). Screening and supporting through schools: educational experiences and needs of adolescents living with HIV in a South African cohort. *BMC Public Health*, *19*(1), N.PAG-N.PAG. https://doi.org/10.1186/s12889-019-6580-0	South Africa	Non-standard use
Trude, A. C. B., Richter, L. M., Behrman, J. R., Stein, A. D., Menezes, A. M. B., & Black, M. M. (2021). Effects of responsive caregiving and learning opportunities during pre-school ages on the association of early adversities and adolescent human capital: an analysis of birth cohorts in two middle-income countries. *The Lancet Child & Adolescent Health*, *5*(1), 37–46. https://doi.org/https://doi.org/10.1016/S2352-4642(20)30309-6	South Africa	Non-standard use
Waller, R., Gardner, F., & Cluver, L. (2014). Shared and unique predictors of antisocial and substance use behavior among a nationally representative sample of South African youth. *Aggression & Violent Behavior*, *19*(6), 629–636. https://doi.org/10.1016/j.avb.2014.09.002	South Africa	Non-standard use
West Africa (*n* = 4)		
Bishop, S. A., Owoloko, E. A., Okagbue, H. I., Oguntunde, P. E., Odetunmibi, O. A., & Opanuga, A. A. (2017). Survey datasets on the externalizing behaviors of primary school pupils and secondary school students in some selected schools in Ogun State, Nigeria. *Data in Brief*, *13*, 469–479. https://doi.org/10.1016/j.dib.2017.06.025	Nigeria	Tool underwent extensive adaptation (e.g., with new items)
Bishop, S. A., Okagbue, H. I., & Odukoya, J. A. (2020). Statistical analysis of childhood and early adolescent externalizing behaviors in a middle low income country. *Heliyon, 6*(2), e03377. doi:10.1016/j.heliyon.2020.e03377	Nigeria	Tool underwent extensive adaptation (e.g., with new items)
Bolton, S., McDonald, D., Curtis, E., Kelly, S., & Gallagher, L. (2014). Autism in a recently arrived immigrant population. *European Journal of Pediatrics*, *173*(3), 337–343. https://doi.org/10.1007/s00431-013-2149-6	Nigeria and Congo	ASEBA data not reported
Roche, K. M., Bingenheimer, J. B., & Ghazarian, S. R. (2016). The dynamic interdependence between family support and depressive symptoms among adolescents in Ghana. *International Journal of Public Health*, *61*(4), 487–494. https://doi.org/10.1007/s00038-015-0781-9	Ghana	Non-standard use
Central Africa (*n* = 2)		
See Bolton et al. (2014)	Congo	
Ibinga, E., Ngoungou, E. B., Olliac, B., Hounsossou, C. H., Dalmay, F., Mouangue, G., Ategbo, S. J., Preux, P.-M., & Druet-Cabanac, M. (2015). Impact of epilepsy on children and parents in Gabon. *Epilepsy & Behavior*, *44*, 110–116. https://doi.org/10.1016/j.yebeh.2014.12.035	Gabon	Non-standard use
Country Unknown (*n* = 20)		
Benenson, J. F., Sinclair, N., & Dolenszky, E. (2006). Children's and adolescents' expectations of aggressive responses to provocation: Females predict more hostile reactions in compatible dyadic relationships. *Social Development*, *15*(1), 65–81. https://doi.org/10.1111/j.1467-9507.2006.00330.x		Country of origin in Africa not specified

Cipolletta, S., Spina, G., & Spoto, A. (2018). Psychosocial functioning, self-image, and quality of life in children and adolescents with neurofibromatosis type 1. *Child: Care, Health & Development*, *44*(2), 260–268. https://doi.org/10.1111/cch.12496 Country of origin in Africa not specified

de la Osa, N., Granero, R., Trepat, E., Domenech, J., & Ezpeleta, L. (2016). The discriminative capacity of CBCL/1½-5-DSM5 scales to identify disruptive and internalizing disorders in preschool children. *European Child & Adolescent Psychiatry*, *25*(1), 17–23. https://doi.org/10.1007/s00787-015-0694-4 Country of origin in Africa not specified

Deater-Deckard, K., & Petrill, S. A. (2004). Parent–child dyadic mutuality and child behavior problems: an investigation of gene–environment processes. *Journal of Child Psychology & Psychiatry*, *45*(6), 1171–1179. https://doi.org/10.1111/j.1469-7610.2004.00309.x Country of origin in Africa not specified

Derluyn, I., & Broekaert, E. (2007). Different perspectives on emotional and behavioural problems in unaccompanied refugee children and adolescents. *Ethnicity & Health*, *12*(2), 141–162. https://doi.org/10.1080/13557850601002296 Country of origin in Africa not specified

Doom, J. R., Georgieff, M. K., & Gunnar, M. R. (2015). Institutional care and iron deficiency increase ADHD symptomology and lower IQ 2.5-5 years post-adoption. *Developmental Science*, *18*(3), 484–494. https://doi.org/10.1111/desc.12223 Country of origin in Africa not specified

Heikkilä, A.-R., Elovainio, M., Raaska, H., Matomäki, J., Sinkkonen, J., & Lapinleimu, H. (2021). Intestinal parasites may be associated with later behavioral problems in internationally adopted children. *PLoS One*, *16*(1), e0245786. https://doi.org/https://doi.org/10.1371/journal.pone.0245786 Country of origin in Africa not specified

Jackson, M. I., Kiernan, K., & McLanahan, S. (2012). Immigrant-native differences in child health: Does maternal education narrow or widen the gap? *Child Development*, *83*(5), 1501–1509. https://doi.org/10.1111/j.1467-8624.2012.01811.x Country of origin in Africa not specified

Leventhal, T., & Shuey, E. A. (2014). Neighborhood context and immigrant young children's development. *Developmental Psychology*, *50*(6), 1771–1787. https://doi.org/10.1037/a0036424 Country of origin in Africa not specified

Löfholm, C. A., Olsson, T., Sundell, K., & Hansson, K. (2009). Multisystemic therapy with conduct-disordered young people: Stability of treatment outcomes two years after intake. *Evidence & Policy: A Journal of Research, Debate and Practice*, *5*(4), 373–397. https://doi.org/https://doi.org/10.1332/174426409x478752 Country of origin in Africa not specified

Meir, Y., Slone, M., & Lavi, I. (2012). Children of illegal migrant workers: Life circumstances and mental health. *Children & Youth Services Review*, *34*(8), 1546–1552. https://doi.org/10.1016/j.childyouth.2012.04.008 Country of origin in Africa not specified

Meir, Y., Slone, M., & Levis, M. (2014). A randomized controlled study of a group intervention program to enhance mental health of children of illegal migrant workers. *Child & Youth Care Forum*, *43*(2), 165–180. https://doi.org/10.1007/s10566-013-9237-7 Country of origin in Africa not specified

Pace, C. S., Di Folco, S., & Guerriero, V. (2018). Late-adoptions in adolescence: Can attachment and emotion regulation influence behaviour problems? A controlled study using a moderation approach. *Clinical Psychology & Psychotherapy*, *25*(2), 250–262. https://doi.org/10.1002/cpp.2158	Country of origin in Africa not specified
Pearl, E. S. (2008). Parent-child interaction therapy with an immigrant family exposed to domestic violence. *Clinical Case Studies*, *7*(1), 25–41. https://doi.org/10.1177/1534650107300939	Country of origin in Africa not specified; case study
Scheper, F. Y., Abrahamse, M. E., Jonkman, C. S., Schuengel, C., Lindauer, R. J. L., De Vries, A. L. C., Doreleijers, T. A. H., & Jansen, L. M. C. (2016). Inhibited attachment behaviour and disinhibited social engagement behaviour as relevant concepts in referred home reared children. *Child: Care, Health and Development*, *42*(4), 544–552. https://doi.org/https://doi.org/10.1111/cch.12319	Country of origin in Africa not specified
Stellern, S., Esposito, E., Mliner, S., Pears, K., & Gunnar, M. (2014). Increased freezing and decreased positive affect in postinstitutionalized children. *Journal of Child Psychology & Psychiatry*, *55*(1), 88–95. https://doi.org/10.1111/jcpp.12123	Country of origin in Africa not specified
Sundell, K., Hansson, K., Löfholm, C. A., Olsson, T., Gustle, L.-H., & Kadesjö, C. (2008). The transportability of multisystemic therapy to Sweden: Short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology*, *22*(4), 550–560. https://doi.org/10.1037/a0012790	Country of origin in Africa not specified
van Ee, E., Kleber, R. J., & Mooren, T. T. M. (2012). War trauma lingers on: Associations between maternal posttraumatic stress disorder, parent-child interaction, and child development. *Infant Mental Health Journal*, *33*(5), 459–468. https://doi.org/10.1002/imhj.21324	Country of origin in Africa not specified
Villabø, M., Gere, M., Torgersen, S., March, J., & Kendall, P. (2012). Diagnostic efficiency of the child and parent versions of the multidimensional anxiety scale for children. *Journal of Clinical Child & Adolescent Psychology*, *41*(1), 75–85. https://doi.org/10.1080/15374416.2012.632350	Country of origin in Africa not specified
Webb, H. J., Thomas, R., McGregor, L., Avdagic, E., & Zimmer-Gembeck, M. J. (2017). An evaluation of parent–child interaction therapy with and without motivational enhancement to reduce attrition. *Journal of Clinical Child & Adolescent Psychology*, *46*(4), 537–550. https://doi.org/10.1080/15374416.2016.1247357	Country of origin in Africa not specified

Appendix 1

Journal articles with ASEBA forms, a sub-Saharan African sample, but no psychometric properties (n = 87)

South Africa (*n* = 41)

Asanbe, C., Moleko, A.-G., Visser, M., Thomas, A., Makwakwa, C., Salgado, W., & Tesnakis, A. (2016). Parental HIV/AIDS and psychological health of younger children in South Africa. *Journal of Child & Adolescent Mental Health*, *28*(2), 175–185. https://doi.org/10.2989/17280583.2016.1216853

Barbarin, O. A., Richter, L., & deWet, T. (2001). Exposure to violence, coping resources, and psychological adjustment of South African children. *The American Journal of Orthopsychiatry*, *71*(1), 16–25. https://doi.org/10.1037/0002-9432.71.1.16

Bell, C. C., Bhana, A., Petersen, I., McKay, M. M., Gibbons, R., Bannon, W., & Amatya, A. (2008). Building protective factors to offset sexually risky behaviors among black youths: a randomized control trial. *Journal of the National Medical Association*, *100*(8), 936–944. https://doi.org/10.1016/s0027-9684(15)31408-5

Bozicevic, L., De Pascalis, L., Schuitmaker, N., Tomlinson, M., Cooper, P. J., & Murray, L. (2016). Longitudinal association between child emotion regulation and aggression, and the role of parenting: A comparison of three cultures. *Psychopathology*, *49*(4), 228–235. https://doi.org/10.1159/000447747

Cluver, C. A., Charles, W., van der Merwe, C., Bezuidenhout, H., Nel, D., Groenewald, C., Brink, L., Hesselman, S., Bergman, L., & Odendaal, H. (2019). The association of prenatal alcohol exposure on the cognitive abilities and behaviour profiles of 4-year-old children: a prospective cohort study. *Bjog*, *126*(13), 1588–1597. https://doi.org/10.1111/1471-0528.15947

Cluver, L. D., Meinck, F., Steinert, J. I., Shenderovich, Y., Doubt, J., Romero, R. H., Lombard, C. J., Redfern, A., Ward, C. L., Tsoanyane, S., Nzima, D., Sibanda, N., Wittesaele, C., De Stone, S., Boyes, M. E., Catanho, R., Lachman, J. M., Salah, N., Nocuza, M., & Gardner, F. (2018). Parenting for lifelong health: A pragmatic cluster randomised controlled trial of a non-commercialised parenting programme for adolescents and their families in South Africa. *BMJ Global Health*, *3*(1). https://doi.org/10.1136/bmjgh-2017-000539

Cluver, L. D., Gardner, F., & Operario, D. (2008). Effects of stigma on the mental health of adolescents orphaned by AIDS. *Journal of Adolescent Health*, *42*(4), 410–417. https://doi.org/10.1016/j.jadohealth.2007.09.022

Cluver, L. D., Orkin, F. M., Meinck, F., Boyes, M. E., Yakubovich, A. R., & Sherr, L. (2016). Can Social Protection Improve Sustainable Development Goals for Adolescent Health? *PLoS One*, *11*(10), 1–20. https://doi.org/10.1371/journal.pone.0164808

Cluver, L. D., Orkin, F. M., Meinck, F., Boyes, M. E., & Sherr, L. (2016). Structural drivers and social protection: mechanisms of HIV risk and HIV prevention for South African adolescents. *Journal of the International AIDS Society*, *19*(1), 20646. https://doi.org/10.7448/ias.19.1.20646

Cluver, L., Gardner, F., & Operario, D. (2007). Psychological distress amongst AIDS-orphaned children in urban South Africa. *Journal of Child Psychology & Psychiatry*, *48*(8), 755–763. https://doi.org/10.1111/j.1469-7610.2007.01757.x

Cluver, L., Gardner, F., & Operario, D. (2009). Poverty and psychological health among AIDS-orphaned children in Cape Town, South Africa. *AIDS Care*, *21*(6), 732–741. https://doi.org/10.1080/09540120802511885

Cluver, L., Meinck, F., Shenderovich, Y., Ward, C. L., Romero, R. H., Redfern, A., Lombard, C., Doubt, J., Steinert, J., Catanho, R., Wittesaele, C., De Stone, S., Salah, N., Mpimpilashe, P., Lachman, J., Loening, H., Gardner, F., Blanc, D., Nocuza, M., & Lechowicz, M. (2016). A parenting programme to prevent abuse of adolescents in South Africa: Study protocol for a randomised controlled trial. *Trials*, *17*(1), 328. https://doi.org/10.1186/s13063-016-1452-8

Cluver, L., Operario, D., & Gardner, F. (2009). Parental illness, caregiving factors and psychological distress among children orphaned by acquired immune deficiency syndrome (AIDS) in South Africa. *Vulnerable Children and Youth Studies*, *4*(3), 185–198. https://doi.org/10.1080/17450120902730196

[a]Cortina, M. A., Stein, A., Kahn, K., Hlungwani, T. M., Holmes, E. A., & Fazel, M. (2016). Cognitive styles and psychological functioning in rural South African school students: Understanding influences for risk and resilience in the face of chronic adversity. *Journal of Adolescence*, *49*, 38–46. https://doi.org/10.1016/j.adolescence.2016.01.010

De Stone, S., F. Meinck, L. Sherr, et al. (2016). Factors associated with good and harsh parenting of pre-adolescents and adolescents in Southern Africa, *Innocenti Working Paper* No. 2016-20, UNICEF Office of Research, Florence.

Dollman, A. K., Figaji, A. A., & Schrieff-Elson, L. E. (2017). Academic and Behavioral Outcomes in School-Age South African Children Following Severe Traumatic Brain Injury. *Frontiers in Neuroanatomy*, *11*. https://doi.org/10.3389/fnana.2017.00121

Donald, K. A. M., Mathema, H., Thomas, K. G. F., & Wilmshurst, J. M. (2011). Intellectual and behavioral functioning in a South African cohort of boys with Duchenne muscular dystrophy. *Journal of Child Neurology*, *26*(8), 963–969. https://doi.org/10.1177/0883073811399149

Ebersöhn, L., Eloff, I., Finestone, M., Grobler, A., & Moen, M. (2015). Telling stories and adding scores: Measuring resilience in young children affected by maternal HIV and AIDS. *African Journal of AIDS Research*, *14*(3), 219–227. https://doi.org/10.2989/16085906.2015.1052822

Eloff, I., Finestone, M., Makin, J. D., Boeving-Allen, A., Visser, M., Ebersöhn, L., Ferreira, R., Sikkema, K. J., Briggs-Gowan, M. J., & Forsyth, B. W. C. (2014). A randomized clinical trial of an intervention to promote resilience in young children of HIV-positive mothers in South Africa. *Aids*, *28*(SUPPL. 3), S347–S357. https://doi.org/10.1097/QAD.0000000000000335

Eloff, I., Finestone, M., & Forsyth, B. (2016). HIV/AIDS Infected Mothers' Experience of a Group Intervention to Enhance Their Children's Behavior. *South African Journal of Education*, *36*(2). http://dx.doi.org/10.15700/saje.v36n2a1285

Garman, E. C., Cois, A., Tomlinson, M., Rotheram-Borus, M. J., & Lund, C. (2019). Course of perinatal depressive symptoms among South African women: associations with child outcomes at 18 and 36 months. *Social Psychiatry and Psychiatric Epidemiology*, 1–13. https://doi.org/10.1007/s00127-019-01665-2

Hoare, J., Phillips, N., Brittain, K., Myer, L., Zar, H. J., & Stein, D. J. (2019). Mental Health and Functional Competence in the Cape Town Adolescent Antiretroviral Cohort. *J Acquir Immune Defic Syndr*, *81*(4), e109–e116. https://doi.org/10.1097/qai.0000000000002068

Hoare, J., Fouche, J.-P., Phillips, N., Joska, J. A., Myer, L., Zar, H. J., & Stein, D. J. (2018). Structural brain changes in perinatally HIV-infected young adolescents in South Africa. *AIDS*, *32*(18), 2707–2718. https://doi.org/10.1097/QAD.0000000000002024

Hoare, J., Phillips, N., Joska, J. A., Paul, R., Donald, K. A., Stein, D. J., & Thomas, K. G. F. (2016). Applying the HIV-associated neurocognitive disorder diagnostic criteria to HIV-infected youth. *Neurology*, *87*(1), 86–93. https://doi.org/10.1212/WNL.0000000000002669

Lanesman, T. H., & Schrieff, L. E. (2020). Implementation of an attention training programme with a sample of children who have sustained traumatic brain injuries in South Africa: A pilot study. *Neuropsychological Rehabilitation*, 1–29. https://doi.org/10.1080/09602011.2020.1782233

Louw, K.-A., Ipser, J., Phillips, N., & Hoare, J. (2016). Correlates of emotional and behavioural problems in children with perinatally acquired HIV in Cape Town, South Africa. *AIDS Care*, *28*(7), 842–850. https://doi.org/10.1080/09540121.2016.1140892

Naicker, N., Richter, L., Mathee, A., Becker, P., & Norris, S. A. (2012). Environmental lead exposure and socio-behavioural adjustment in the early teens: the birth to twenty cohort. *Science of the Total Environment*, *414*, 120–125. https://doi.org/10.1016/j.scitotenv.2011.11.013

Nothling, J., Martin, C. L., Laughton, B., Cotton, M. F., & Seedat, S. (2013). Maternal post-traumatic stress disorder, depression and alcohol dependence and child behaviour outcomes in mother-child dyads infected with HIV: a longitudinal study. *BMJ Open*, *3*(12), e003638. https://doi.org/10.1136/bmjopen-2013-003638

[b]Phillips, N., Thomas, K. G. F., Mtukushe, B., Myer, L., Zar, H. J., Stein, D. J., & Hoare, J. (2021). Youth perinatal HIV-associated neurocognitive disorders: association with functional impairment. *AIDS Care*, 1–5. https://doi.org/10.1080/09540121.2021.1891191

Rochat, T. J., Houle, B., Stein, A., Mitchell, J., & Bland, R. M. (2019). Maternal alcohol use and children's emotional and cognitive outcomes in rural South Africa. *SAMJ: South African Medical Journal*, *109*(7), 526–534. https://doi.org/10.7196/samj.2019.v109i7.13120

Rochat, T., Stein, A., Cortina-Borja, M., Tanser, F., & Bland, R. M. (2017). The Amagugu intervention to increase disclosure of maternal HIV to HIV-uninfected primary-school aged children in Southern Africa: A randomised controlled trial. *Lancet HIV*, *4*(12), e566–e576. https://doi.org/10.1016/s2352-3018(17)30133-9

Rochat, T. J., Arteche, A. X., Stein, A., Mkwanazi, N., & Bland, R. M. (2014). Maternal HIV disclosure to young HIV-uninfected children: an evaluation of a family-centred intervention in South Africa. *AIDS (London, England)*, *28 Suppl 3*, S331-41. https://doi.org/10.1097/QAD.0000000000000333

Rochat, T. J., Houle, B., Stein, A., Pearson, R. M., & Bland, R. M. (2018). Prevalence and risk factors for child mental disorders in a population-based cohort of HIV-exposed and unexposed African children aged 7–11 years. *European Child & Adolescent Psychiatry*, *27*(12), 1607–1620. https://doi.org/10.1007/s00787-018-1146-8

Rossouw, J., Yadin, E., Alexander, D., & Seedat, S. (2018). Prolonged exposure therapy and supportive counselling for posttraumatic stress disorder in adolescents in a community-based sample, including experiences of stakeholders: study protocol for a comparative randomized controlled trial using task-shifting. *BMC Psychiatry*, *18*(1), N.PAG-N.PAG. https://doi.org/10.1186/s12888-018-1873-x

Rotheram-Borus, M. J., Arfer, K. B., Christodoulou, J., Comulada, W. S., Stewart, J., Tubert, J. E., & Tomlinson, M. (2019). The association of maternal alcohol use and paraprofessional home visiting with children's health: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *87*(6), 551–562. https://doi.org/10.1037/ccp0000408

Rotheram-Fuller, E. J., Tomlinson, M., Scheffler, A., Weichle, T. W., Hayati Rezvan, P., Comulada, W. S., & Rotheram-Borus, M. J. (2018). Maternal patterns of antenatal and postnatal depressed mood and the impact on child health at 3-years postpartum. *Journal of Consulting and Clinical Psychology*, *86*(3), 218–230. https://doi.org/10.1037/ccp0000281

Sabet, F., Richter, L. M., Ramchandani, P. G., Stein, A., Quigley, M. A., & Norris, S. A. (2009). Low birthweight and subsequent emotional and behavioural outcomes in 12-year-old children in Soweto, South Africa: Findings from Birth to Twenty. *International Journal of Epidemiology*, *38*(4), 944–954. https://doi.org/10.1093/ije/dyp204

St Clair, M. C., Skeen, S., Marlow, M., & Tomlinson, M. (2019). Relationships between concurrent language ability and mental health outcomes in a South African sample of 13-year-olds. *PLoS One*, *14*(9), e0221242. https://doi.org/10.1371/journal.pone.0221242

van Dyk, J., Ramanjam, V., Church, P., Koren, G., & Donald, K. (2014). Maternal methamphetamine use in pregnancy and long-term neurodevelopmental and behavioral deficits in children. *Journal of Population Therapeutics and Clinical Pharmacology*, *21*(2), e185–e196.

Wait, J. W. V. (2002). Tuberculosis Meningitis and Attention Deficit Hyperactivity Disorder in Children. *Journal of Tropical Pediatrics*, *48*(5), 294–299. https://doi.org/10.1093/tropej/48.5.294

Wait, J. W., & Schoeman, J. F. (2010). Behaviour Profiles After Tuberculous Meningitis. *Journal of Tropical Pediatrics*, *56*(3), 166–171. https://doi.org/10.1093/tropej/fmp080

[a]Internal consistency analyses were conducted but not reported.
[b]Psychometric analyses were performed but they were not conventional.

Kenya (*n* = 18)

Denckla, C. A., Ndetei, D. M., Mutiso, V. N., Musyimi, C. W., Musau, A. M., Nandoya, E. S., Anderson, K. K., Milanovic, S., Henderson, D., & McKenzie, K. (2017). Psychometric properties of the Ndetei-Othieno-Kathuku (NOK) Scale: A mental health assessment tool for an African setting. *J Child Adolesc Ment Health*, *29*(1), 39–49. https://doi.org/10.2989/17280583.2017.1310729

Di Giunta, L., Rothenberg, W. A., Lunetti, C., Lansford, J. E., Pastorelli, C., Eisenberg, N., Thartori, E., Basili, E., Favini, A., & Yotanyamaneewong, S. (2020). Longitudinal associations between mothers' and fathers' anger/irritability expressiveness, harsh parenting, and adolescents' socioemotional functioning in nine countries. *Developmental Psychology*, *56*(3), 458. https://doi.org/10.1037/dev0000849

Dodge, K. A., Malone, P. S., Lansford, J. E., Sorbring, E., Skinner, A. T., Tapanya, S., Uribe Tiradod, L. M., Zelli, A., Alampay, L. P., Al-Hassan, S. M., Bacchini, D., Bombi, A. S., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Oburu, P., & Pastorelli, C. (2015). Hostile attributional bias and aggressive behavior in global context. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *112*(30), 9310–9315. https://doi.org/10.1073/pnas.1418572112

Karanja, S. W., Kiburi, S. K., Kang'ethe, R., & Othieno, C. J. (2021). Emotional and behavioral problems in children with epilepsy attending the pediatric neurology clinic at a referral hospital in Kenya. *Epilepsy Behav*, *114*(Pt A), 107477. https://doi.org/10.1016/j.yebeh.2020.107477

Kariuki, S. M., Abubakar, A., Kombe, M., Kazungu, M., Odhiambo, R., Stein, A., & Newton, C. R. J. C. (2017). Burden, risk factors, and comorbidities of behavioural and emotional problems in Kenyan children: a population-based study. *The Lancet Psychiatry*, *4*(2), 136–145. https://doi.org/10.1016/S2215-0366(16)30403-5

Kariuki, S. M., Abubakar, A., Kombe, M., Kazungu, M., Odhiambo, R., Stein, A., & Newton, C. R. J. C. (2018). Prevalence, risk factors and behavioural and emotional comorbidity of acute seizures in young Kenyan children: a population-based study. *BMC Medicine*, *16*(1), 35. https://doi.org/10.1186/s12916-018-1021-y

Lansford, J. E., Godwin, J., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Chen, B.-B., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Steinberg, L., Tapanya, S., Alampay, L. P., Uribe Tirado, L. M., & Zelli, A. (2018). Longitudinal associations between parenting and youth adjustment in twelve cultural groups: Cultural normativeness of parenting as a moderator. *Developmental Psychology*, *54*(2), 362–377. https://doi.org/10.1037/dev0000416

Lansford, J. E., Sharma, C., Malone, P. S., Woodlief, D., Dodge, K. A., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Tirado, L. M. U., Zelli, A., Al-Hassan, S. M., Alampay, L. P., Bacchini, D., Bombi, A. S., Bornstein, M. H., Chang, L., Deater-Deckard, K., & Di Giunta, L. (2014). Corporal punishment, maternal warmth, and child adjustment: A longitudinal study in eight countries. *Journal of Clinical Child and Adolescent Psychology*, *43*(4), 670–685. https://doi.org/10.1080/15374416.2014.893518

Magai, D. N., & Koot, H. M. (2019). Quality of life in children and adolescents in Central Kenya: associations with emotional and behavioral problems. *Quality of Life Research*, *28*(5), 1271–1279. https://doi.org/10.1007/s11136-019-02099-8

Mutiso, V., Musyimi, C., Tele, A., Gitonga, I., & Ndetei, D. (2020). Feasibility study on the mhGAP-IG as a tool to enhance parental awareness of symptoms of mental disorders in lower primary (6-10 year old) school-going children: Towards inclusive child mental health services in a Kenyan setting. *Early Interv Psychiatry*. https://doi.org/10.1111/eip.12963

Mutiso, V., Tele, A., Musyimi, C., Gitonga, I., Musau, A., & Ndetei, D. (2018). Effectiveness of life skills education and psychoeducation on emotional and behavioral problems among adolescents in institutional care in Kenya: a longitudinal study. *Child and Adolescent Mental Health*, *23*(4), 351–358. https://doi.org/10.1111/camh.12232

Mutiso, V. N., Musyimi, C. W., Tele, A., & Ndetei, D. M. (2017). Epidemiological patterns and correlates of mental disorders among orphans and vulnerable children under institutional care. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 65–75. https://doi.org/10.1007/s00127-016-1291-7

Ndetei, D. M., Mutiso, V., Gitonga, I., Agudile, E., Tele, A., Birech, L., Musyimi, C., & McKenzie, K. (2019). World Health Organization life-skills training is efficacious in reducing youth self-report scores in primary school going children in Kenya. *Early Intervention in Psychiatry*, *13*(5), 1146–1154. https://doi.org/10.1111/eip.12745

Ndetei, D., Mutiso, V., Maraj, A., Anderson, K., Musyimi, C., Musau, A., Tele, A., Gitonga, I., & McKenzie, K. (2019). Towards Understanding the Relationship Between Psychosocial Factors and Ego Resilience Among Primary School Children in a Kenyan Setting: A Pilot Feasibility Study. *Community Mental Health Journal*, *55*(6), 1038–1046. https://doi.org/10.1007/s10597-019-00425-5

Ndetei, D., Mutiso, V., Musyimi, C., Mokaya, A., Anderson, K., McKenzie, K., Musau, A., Ndetei, D. M., Mokaya, A. G., & Anderson, K. K. (2016). The prevalence of mental disorders among upper primary school children in Kenya. *Social Psychiatry & Psychiatric Epidemiology*, *51*(1), 63–71. https://doi.org/10.1007/s00127-015-1132-0

Rothenberg, W. A., Lansford, J. E., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., & Malone, P. S. (2020). Examining effects of parent warmth and control on internalizing behavior clusters from age 8 to 12 in 12 cultural groups in nine countries. *Journal of Child Psychology and Psychiatry*, *61*(4), 436–446. https://doi.org/10.1111/jcpp.13138

Schenck-Fontaine, A., Lansford, J. E., Skinner, A. T., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Oburu, P., Pastorelli, C., Sorbring, E., & Steinberg, L. (2020). Associations between perceived material deprivation, parents' discipline practices, and children's behavior problems: An international perspective. *Child Development*, *91*(1), 307–326. https://doi.org/10.1111/cdev.13151

Weisz, J. R., Sigman, M., Weiss, B., & Mosk, J. (1993). Parent reports of behavioral and emotional problems among children in Kenya, Thailand, and the United States. *Child Dev*, *64*(1), 98–109. https://doi.org/10.1111/j.1467-8624.1993.tb02897.x

Uganda (*n* = 13)

Bangirana, P., Allebeck, P., Boivin, M. J., John, C. C., Page, C., Ehnvall, A., & Musisi, S. (2011). Cognition, behaviour and academic skills after cognitive rehabilitation in Ugandan children surviving severe malaria: a randomised trial. *BMC Neurology*, *11*(1), 96. https://doi.org/10.1186/1471-2377-11-96

Bangirana, P., Giordani, B., John, C. C., Page, C., Opoka, R. O., & Boivin, M. J. (2009). Immediate Neuropsychological and Behavioral Benefits of Computerized Cognitive Rehabilitation in Ugandan Pediatric Cerebral Malaria Survivors. *Journal of Developmental & Behavioral Pediatrics*, *30*(4), 310–318. https://doi.org/10.1097/DBP.0b013e3181b0f01b

Bangirana, P., Musisi, S., Boivin, M. J., Ehnvall, A., John, C. C., Bergemann, T. L., & Allebeck, P. (2011). Malaria with neurological involvement in Ugandan children: effect on cognitive ability, academic achievement and behaviour. *Malaria Journal*, *10*(1), 334. https://doi.org/10.1186/1475-2875-10-334

Boivin, M. J., Sikorskii, A., Nakasujja, N., Ruisenor-Escudero, H., Familiar-Lopez, I., Opoka, R. O., & Giordani, B. (2019). Evaluating Immunopathogenic Biomarkers During Severe Malaria Illness as Modifiers of the Neuropsychologic Benefits of Computer Cognitive Games Rehabilitation in Ugandan Children. *Pediatr Infect Dis J*, *38*(8), 840–848. https://doi.org/10.1097/inf.0000000000002367

Boivin, M. J., Bangirana, P., Nakasujja, N., Page, C. F., Shohet, C., Givon, D., Bass, J. K., Opoka, R. O., & Klein, P. S. (2013). A Year-Long Caregiver Training Program Improves Cognition in Preschool Ugandan Children with Human Immunodeficiency Virus. *The Journal of Pediatrics*, *163*(5), 1409-1416.e5. https://doi.org/10.1016/j.jpeds.2013.06.055

Boivin, M. J., Bangirana, P., Nakasujja, N., Page, C. F., Shohet, C., Givon, D., Bass, J. K., Opoka, R. O., & Klein, P. S. (2013). A Year-long Caregiver Training Program to Improve Neurocognition in Preschool Ugandan HIV-exposed Children. *Journal of Developmental & Behavioral Pediatrics*, *34*(4), 269–278. https://doi.org/10.1097/DBP.0b013e318285fba9

Boivin, M. J., Nakasujja, N., Sikorskii, A., Opoka, R. O., & Giordani, B. (2016). A Randomized Controlled Trial to Evaluate if Computerized Cognitive Rehabilitation Improves Neurocognition in Ugandan Children with HIV. *AIDS Research and Human Retroviruses*, *32*(8), 743–755. https://doi.org/10.1089/aid.2016.0026

Boivin, M. J., Nakasujja, N., Sikorskii, A., Ruiseñor-Escudero, H., Familiar-Lopez, I., Walhof, K., van der Lugt, E. M., Opoka, R. O., & Giordani, B. (2019). Neuropsychological benefits of computerized cognitive rehabilitation training in Ugandan children surviving severe malaria: A randomized controlled trial. *Brain Research Bulletin*, *145*, 117–128. https://doi.org/10.1016/j.brainresbull.2018.03.002

Hickson, M. R., Conroy, A. L., Bangirana, P., Opoka, R. O., Idro, R., Ssenkusu, J. M., & John, C. C. (2019). Acute kidney injury in Ugandan children with severe malaria is associated with long-term behavioral problems. *PLoS One*, *14*(12), e0226405. https://doi.org/10.1371/journal.pone.0226405

Ojiambo, D., & Bratton, S. C. (2014). Effects of group activity play therapy on problem behaviors of preadolescent Ugandan orphans. *Journal of Counseling & Development*, *92*(3), 355–365. https://doi.org/10.1002/j.1556-6676.2014.00163.x

Okello, J., Nakimuli-Mpungu, E., Musisi, S., Broekaert, E., & Derluyn, I. (2013). War-related trauma exposure and multiple risk behaviors among school-going adolescents in Northern Uganda: the mediating role of depression symptoms. *J Affect Disord*, *151*(2), 715–721. https://doi.org/10.1016/j.jad.2013.07.030

Ssemata, A. S., Hickson, M., Ssenkusu, J. M., Cusick, S. E., Nakasujja, N., Opoka, R. O., Kroupina, M., Georgieff, M. K., Bangirana, P., & John, C. C. (2020). Delayed iron does not alter cognition or behavior among children with severe malaria and iron deficiency. *Pediatric Research*, *88*(3), 429–437. https://doi.org/10.1038/s41390-020-0957-8

Ssenkusu, J. M., Hodges, J. S., Opoka, R. O., Idro, R., Shapiro, E., John, C. C., & Bangirana, P. (2016). Long-term Behavioral Problems in Children With Severe Malaria. *Pediatrics*, *138*(5). https://doi.org/10.1542/peds.2016-1965

Ethiopia (*n* = 4)

Rescorla, L., Achenbach, T. M., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., Bird, H., Broberg, A., Dobrean, A., Döpfner, M., Erol, N., Forns, M., Hannesdottir, H., Kanbayashi, Y., Lambert, M. C., Leung, P., Minaei, A., Mulatu, M. S., Novik, T. S., … Verhulst, F. (2007). Epidemiological comparisons of problems and positive qualities reported by adolescents in 24 countries. *Journal of Consulting and Clinical Psychology*, *75*(2), 351–358. https://doi.org/10.1037/0022-006X.75.2.351

Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., Bird, H., Chen, W., Dobrean, A., Döpfner, M., Erol, N., Fombonne, E., Fonseca, A., Frigerio, A., Grietens, H., Hannesdottir, H., Kanbayashi, Y., Lambert, M., Larsson, B., … Verhulst, F. (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. https://doi.org/10.1177/10634266070150030101

Rescorla, L., Ivanova, M. Y., Achenbach, T. M., Begovac, I., Chahed, M., Drugli, M. B., Emerich, D. R., Fung, D. S. S., Haider, M., Hansson, K., Hewitt, N., Jaimes, S., Larsson, B., Maggiolini, A., Marković, J., Mitrović, D., Moreira, P., Oliveira, J. T., Olsson, M., … Zhang, E. Y. (2012). International epidemiology of child and adolescent psychopathology II: Integration and applications of dimensional findings from 44 societies. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(12), 1273–1283. https://doi.org/10.1016/j.jaac.2012.09.012

Tadesse, A. W., Berhane Tsehay, Y., Girma Belaineh, B., & Alemu, Y. B. (2012). Behavioral and emotional problems among children aged 6–14 years on highly active antiretroviral therapy in Addis Ababa: A cross-sectional study. *AIDS Care*, *24*(11), 1359–1367. https://doi.org/10.1080/09540121.2011.650677

Malawi (*n* = 4)

Boivin, M. J., Gladstone, M. J., Vokhiwa, M., Birbeck, G. L., Magen, J. G., Page, C., Semrud-Clikeman, M., Kauye, F., & Taylor, T. E. (2011). Developmental outcomes in Malawian children with retinopathy-positive cerebral malaria. *Tropical Medicine & International Health*, *16*(3), 263–271. https://doi.org/10.1111/j.1365-3156.2010.02704.x

Boivin, M. J., Mohanty, A., Sikorskii, A., Vokhiwa, M., Magen, J. G., & Gladstone, M. (2019). Early and middle childhood developmental, cognitive, and psychiatric outcomes of Malawian children affected by retinopathy positive cerebral malaria. *Child Neuropsychology*, *25*(1), 81–102. https://doi.org/10.1080/09297049.2018.1451497

Boivin, M. J., Vokhiwa, M., Sikorskii, A., Magen, J. G., & Beare, N. A. V. (2014). Cerebral Malaria Retinopathy Predictors of Persisting Neurocognitive Outcomes in Malawian Children. *The Pediatric Infectious Disease Journal*, *33*(8), 821–824. https://doi.org/10.1097/INF.0000000000000296

Madhlopa, Y., Qin, J., & Chen, C. (2020). The relationships between child maltreatment and child behavior problems. Comparative study of Malawi and China. *Children and Youth Services Review*, *119*, 105533. https://doi.org/10.1016/j.childyouth.2020.105533

Democratic Republic of the Congo (*n* = 1)

Matonda-ma-Nzuzi, T., Mampunza Ma Miezi, S., Mpembi, M. N., Mvumbi, D. M., Aloni, M. N., Malendakana, F., Mpaka Mbeya, D., Lelo, G. M., & Charlier-Mikolajczak, D. (2018). Factors associated with behavioral problems and cognitive impairment in children with epilepsy of Kinshasa, Democratic Republic of the Congo. *Epilepsy & Behavior*, *78*, 78–83. https://doi.org/10.1016/j.yebeh.2017.08.030

Ghana (*n* = 1)

Kusi-Mensah, K., Donnir, G., Wemakor, S., Owusu-Antwi, R., & Omigbodun, O. (2019). Prevalence and patterns of mental disorders among primary school age children in Ghana: correlates with academic achievement. *J Child Adolesc Ment Health*, *31*(3), 214–223. https://doi.org/10.2989/17280583.2019.1678477

Tanzania (*n* = 1)

Van de Wijer, L., Mchaile, D. N., de Mast, Q., Mmbaga, B. T., Rommelse, N. N. J., Duinmaijer, A., van der Ven, A. J. A. M., Schellekens, A. F. A., & Kinabo, G. D. (2019). Neuropsychiatric symptoms in Tanzanian HIV-infected children receiving long-term efavirenz treatment: a multicentre, cross-sectional, observational study. *The Lancet HIV*, *6*(4), e250–e258. https://doi.org/10.1016/S2352-3018(18)30329-1

Immigrant samples (*n* = 4)

Huemer, J., Karnik, N., Voelkl-Kernstock, S., Granditsch, E., Plattner, B., Friedrich, M., & Steiner, H. (2011). Psychopathology in African Unaccompanied Refugee Minors in Austria. *Child Psychiatry & Human Development*, *42*(3), 307–319. https://doi.org/10.1007/s10578-011-0219-4

Huemer, J., Völkl-Kernstock, S., Karnik, N., G. Denny, K., Granditsch, E., Mitterer, M., Humphreys, K., Plattner, B., Friedrich, M., Shaw, R. J., & Steiner, H. (2013). Personality and Psychopathology in African Unaccompanied Refugee Minors: Repression, Resilience and Vulnerability. *Child Psychiatry & Human Development*, *44*(1), 39–50. https://doi.org/10.1007/s10578-012-0308-z

Sourander, A. (1998). Behavior Problems and Traumatic Events of Unaccompanied Refugee Minors. *Child Abuse & Neglect*, *22*(7), 719–727. https://doi.org/https://doi.org/10.1016/S0145-2134(98)00053-2

Betancourt, T. S., Berent, J. M., Freeman, J., Frounfelker, R. L., Brennan, R. T., Abdi, S., Maalim, A., Abdi, A., Mishra, T., & Gautam, B. (2020). Family-based mental health promotion for Somali bantu and Bhutanese refugees: feasibility and acceptability trial. *Journal of Adolescent Health*, *66*(3), 336–344. https://doi.org/10.1016/j.jadohealth.2019.08.023

In the following chapter, I provide greater detail on the methods of the study investigating the validity of the Swanson, Nolan, and Pelham ADHD Rating Scale (SNAP-IV) in children with neurodevelopmental disorders. This chapter supplements the methods section in the manuscript that was condensed due to the journal's word count restrictions. I also contextualise this study within its larger umbrella study, and describe the statistical techniques employed for the psychometric analyses.

CHAPTER 3

Methods

This chapter details the methods used in the study "Validity of the SNAP-IV for ADHD assessment in South African children with neurodevelopmental disorders", presented in Chapter 4.

Design and Setting

The study was embedded in a larger project titled "Genetic Characterization of Neurodevelopmental Disorders in South African Populations: The NeuroDEV Study" (https://www.neurodevproject.org/). Current evidence on the genetics and phenotypic characteristics of neurodevelopmental disorders (NDDs) in African populations is limited (Abubakar et al., 2016; Baxter et al., 2015; Malcolm-Smith et al., 2013). The NeuroDEV Study aims to expand our knowledge of the genetics of African children with NDDs (de Menil et al., 2019).

The NeuroDEV study is collecting genetic and in-depth phenotypic data from families in the Western Cape of South Africa and in Kilifi, Kenya. The cohort includes children with NDDs aged 2-17 years. The study has an enrolment target of 3000 participants in South Africa and 2600 participants in Kenya with approximately one-third child cases, one-third child controls, and one-third parents (of case children). Red Cross War Memorial Children's Hospital (RCWMCH) and Tygerberg Hospital (TH) are the primary sites for recruitment and data collection in the Western Cape. More information about the NeuroDEV study is presented in an article by de Menil and colleagues (2019). I also provide relevant details concerning the methods of the umbrella study in the sections below.

A correlational research design was appropriate for the study as it sought to determine the nature of associations between scores of various behavioural tools within a specific sample.

Participants

Eligibility and Exclusion Criteria

Both parents and children were enrolled as study participants for the NeuroDEV South Africa Study. However, this sub-study was only concerned with parent participants. To be eligible for inclusion in this sub-study, parents had to have a child with a confirmed clinical diagnosis of one or more NDD, as assessed by the child's treating clinician. NDDs included Attention-Deficit/Hyperactivity Disorder (ADHD), Autism Spectrum Disorder (ASD), Global Developmental Delay (GDD), Intellectual Disability (ID), Specific Learning Disorders (SLD), or Communication Disorder (CD), as defined by the fifth edition of the

Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013). Key characteristics of these neurodevelopmental disorders are briefly outlined in the Glossary of Terms on page viii. Parents with children who had secondary diagnoses of an NDD were also eligible to participate. Parents were excluded from the study if their child was under the age of six years or over the age of 17 years, had a primary neuro-motor disability (e.g., cerebral palsy), was an inpatient, or was unable to participate in the study. Reasons for being unable to participate included having an acute health condition with intrusive symptoms. The lower age limit was implemented as the SNAP-IV was only administered to parents of children aged six years or older in the NeuroDEV study.

Measures

All measures used in this sub-study were self-report and were administered verbally by the data collection team. No specific training was required to administer any of the measures.

Demographics Questionnaire

The demographics questionnaire (Appendix A) included questions about home language. Administration took approximately five minutes.

Asset Index (AI)

The AI (Appendix B) is a sociodemographic questionnaire including questions about parental educational attainment, employment, individual and family income, housing, and access to services (e.g., running water, electricity, etc). The AI took approximately five minutes to administer. The AI is an adaptation from previous South African studies, including the South African Stress and Health study (Myer et al., 2008) and the Drakenstein Child Health Study (DCHS; Stein et al., 2015). We obtained Afrikaans and Xhosa translations of the AI from the DCHS team. Two Afrikaans- and two Xhosa-speaking members of the NeuroDEV data collection team (myself included) reviewed the Afrikaans and Xhosa translations respectively to check for conceptual equivalence to the original English AI. No modifications were necessary.

Neuromedical Assessment

The neuromedical assessment comprised questions about the child's neuromedical and developmental history, including their diagnosis, current level of spoken language, and level of education. A medical officer administered the tool to the child's parent and extracted information about the child's diagnosis from their medical records. Administration of the questions relevant to this sub-study took approximately five minutes. An abbreviated version

of the neuromedical assessment (i.e., containing only those questions relevant to this sub-study) is presented as Appendix C.

Swanson, Nolan, and Pelham ADHD Rating Scale (4th edition; SNAP-IV)

The SNAP-IV (Appendix D) is an 18-item self-report behavioural rating scale designed to evaluate symptoms of ADHD in children and adolescents (Swanson et al., 2001, 2012). The SNAP-IV comprises two subscales, Inattention and Hyperactivity-Impulsivity, each consisting of nine items corresponding to the DSM-5 ADHD symptom groupings of inattention and hyperactivity-impulsivity respectively (American Psychiatric Association, 2013). Each item is presented as a question (e.g., "How often is your child forgetful in daily activities?"). Parents responded to each item using a four-point Likert scale, ranging from 0 ("Not at all") to 3 ("Often"). Subscale scores were calculated by averaging the responses across the relevant items (i.e., each subscale had a possible score range of 0-3). Details regarding the translation process of the SNAP-IV from English to Afrikaans and Xhosa for this study are presented in Chapter 4.

The SNAP-IV has been used and validated in a few different cultural settings (Amazue et al., 2015; Bussing et al., 2008; Costa et al., 2019; Gau et al., 2008; Grañana et al., 2011; Hall et al., 2020; Ullebø et al., 2012). The SNAP-IV has not been validated for South African children, although it has been used in South Africa for clinical purposes and in one published study with HIV-infected children (Zeegers et al., 2010). One unpublished South African study administered the SNAP-IV to children diagnosed with ADHD and reported excellent internal consistency (Fouché, 2017). Another South African study confirmed the expected two-factor structure (i.e., Inattention and Hyperactivity-Impulsivity) of an earlier version of the SNAP-IV, the Disruptive Behaviour Disorders Rating Scale, across six language groups (Meyer et al., 2004).

Child Behaviour Checklist for Ages 6-18 (CBCL/6-18)

The CBCL/6-18 is a parent-report questionnaire designed to assess specific problematic behaviours, thoughts, and emotions in school-aged children (Achenbach & Rescorla, 2001). It is considered a "gold standard" tool for assessing problematic behaviours in children (Ivanova et al., 2007). The structure of the CBCL/6-18 has already been described in Chapter 2. The study administered the Attention Problems syndrome scale (Appendix E) and the ADHD DSM-oriented scale (Appendix F) to measure ADHD symptoms. The Aggressive Behaviour (Appendix G) and the Withdrawn/Depressed (Appendix H) syndrome scales were used to measure externalising and internalising behaviours respectively. Each item was presented as a statement describing a behaviour (e.g. "Daydreams or gets lost in

thoughts"). For each statement, parents were asked to indicate whether the statement was not true (scored as "0"), sometimes/somewhat true (scored as "1"), or very/often true (scored as "2) of their child currently or within the past six months. Subscale scores were calculated by summing the scores of all relevant items. The full CBCL/6-18 was administered for the NeuroDEV Study, which took approximately 30-40 minutes to administer.

The NeuroDEV principal investigators obtained licenses from the test developer, Thomas M. Achenbach, to administer and score the CBCL/6-18 (License Number #1583-10-26-17). We also obtained Afrikaans and Xhosa translated versions of the CBCL/6-18 from the test developer (License Numbers #649-12-21-11 and #303-8-01-08 for the Afrikaans and Xhosa versions respectively).

A number of local studies have used the CBCL for school-aged children. One study administered the Aggressive Behaviour syndrome scale to 30 Xhosa-speaking parents and estimated good internal consistency of the subscale scores, as measured by coefficient alpha ($\alpha = 0.85$; Cluver et al., 2016). Another study based in Tshwane administered the ADHD DSM-oriented scale to 167 Pedi-, Tswana-, Sotho- and Zulu-speaking parents, and also estimated adequate internal consistency of the subscale scores ($\alpha = 0.74$; Visser et al., 2018).

Procedure

The NeuroDEV data collection team consisted of paediatricians, medical officers, research nurses, psychologists, and research assistants. Two postgraduate students (including myself) joined the core team on different days of the week. The procedure for the NeuroDEV Study included the neuromedical assessment, blood draws from the parents and children, and administration of questionnaires and cognitive assessments. I administered and scored the tools relevant to this sub-study as part of my data collection duties. I collected data for the sub-study between August 2018 and December 2019.

Recruitment

A member of the data collection team typically approached parents in the waiting area of the neurodevelopment clinic at RCWMCH, provided general information about the aims and procedures of the study, and asked parents if they would be interested in participating. If parents expressed interest, a suitable date and time for data collection was arranged. We also recruited participants via referrals from clinicians working at neurology and genetic clinics at RCWMCH and TH, and at special needs schools. More information on the sampling methods for this sub-study within the context of the NeuroDEV Study is presented in Chapter 4.

Data Collection

At the study site, a member of the research team escorted participants to a private consulting room. A research nurse guided the participants through the consent processes in their home language. After obtaining informed consent from participants, all questionnaires were administered in the participant's home (or otherwise preferred) language by a team member fluent in that language. For each family, participation in the NeuroDEV Study took approximately three hours. Administration of the assessments relevant to this sub-study took approximately 45 minutes. Upon conclusion of the session, participants were reimbursed with R250 per family for their time and to cover transportation costs. In addition, a meal was provided for the family and children received a goodie bag with snacks and a toy. If participants enrolled in the study but were unwilling or unable to continue participating at any point during data collection, they were still reimbursed with R250 per family.

Data Management and Storage

All questionnaire responses were directly entered onto secure and encrypted electronic tablets, using REDCap, a data capture system. The tablets were operated by the research team only. Electronic data were stored locally and then uploaded to a secure cloud storage system once a secure internet connection was available. I obtained access to a password-protected file containing data relevant for this sub-study from the NeuroDEV South Africa Project Manager.

Statistical Analyses

In this section, I briefly describe the statistical techniques used in the present study, including confirmatory factor analysis (CFA) to investigate structural validity of the SNAP-IV, correlation analyses to evaluate the construct validity of the SNAP-IV, and ordinal coefficient alpha as a measure of internal consistency. I performed all statistical analyses in RStudio (Version 1.3.1093) for R (Version 4.0.2; R Core Team, 2019).

Descriptive Statistics

First, I tidied and wrangled the raw data using the 'Tidyverse' collection of R packages (Wickham et al., 2019). This process included checking for missing data and impossible values, as well as formatting the raw data for the analyses. I then conducted exploratory data analyses using the 'psych' package (Revelle, 2020). Specifically, I calculated descriptive statistics (means, standard deviations, medians, and frequencies for ordinal variables). Finally, I created histograms and boxplots to visualise the distribution of the data and to check for any outlying data points.

Confirmatory Factor Analysis (CFA)

CFA is concerned with the relationships between observed and latent variables (Kline, 2016). Observed variables are actual observations recorded and collected as data. These correspond to the responses to items on a tool. Latent variables, on the other hand, are hypothetical constructs that cannot be observed or measured directly. Inattention and hyperactivity-impulsivity are examples of latent variables, which, in turn, measure a higher-order latent variable, ADHD. A latent variable can be understood as one continuous dimension along which individuals can vary. Observed variables ('indicators') indirectly measure latent variables ('factors').

CFA confirms an a priori specification (i.e., hypothesis) about the underlying structure of a tool (Rust & Golombok, 2009). In other words, CFA asks whether a tool produces scores that seem to measure the intended underlying construct, and whether the items intended to measure a given underlying construct are actually and exclusively measuring that construct. These analyses require specific hypotheses with regards to the number of underlying constructs (factors), which items (indicators) correspond to which constructs, and whether the constructs covary.

The fundamental principle of CFA is the partitioning of variance (Kline, 2016). Variance is characterized as 'systematic' or 'non-systematic' (i.e., random error). Common variance is shared among the indicators and is presumed to be due to the factor (i.e., it is systematic). For example, we expect SNAP-IV items "forgets things" and "can't pay attention to detail" to share variance (i.e., correlate), because they are both presumed to measure inattention. Unique variance is also systematic but is not explained by the factor. For example, characteristics of the individual indicators (i.e., items) or measurement methods may account for systematic non-shared variance.

The objective of CFA is to estimate the parameters of the measurement model (i.e., factor loadings, variances and covariances, and error covariances, if applicable) to produce a predicted variance–covariance matrix that is as close as possible to the sample variance–covariance matrix (Brown, 2015). A 'fitting' function minimizes the differences between the matrices. For non-normally distributed and polytomous data (i.e., ordinal indicators, such as the SNAP-IV items) robust weighted least squares estimators are appropriate (Flora & Curran, 2004). In the present analysis, I used a weighted least square mean- and variance-adjusted (WLSMV) estimator (Li, 2016). Pearson correlation coefficients may underestimate or distort the relationship between latent variables measured with responses to Likert-type scales, especially if the distributions of observed responses on the Likert scale are skewed

(Kline, 2020). Hence, I used polychoric correlation matrices, which more accurately estimate the relationships between ordinal variables (Brown, 2015).

A chi square (χ^2) test is a text of exact fit between the model (i.e., the predicted variance-covariance matrix) and the data (i.e., the sample variance-covariance matrix). A statistically significant χ^2 supports the alternative hypothesis that the model estimates do not sufficiently reproduce the sample variances and covariances (i.e., the fit between the model and the data is poor). Hence, a non-significant χ^2 test indicates exact fit. Another index of "absolute fit" is the standardized root mean square residual (SRMR). The SRMR is the standardised mean difference between the sample correlation matrix and the model-predicted correlation matrix. Hence, it is derived from a residual correlation matrix (Thompson, 2004). In general, SRMR values ≥ 0.08 may suggest a poorly fitting model (SRMR = 0 indicates 'perfect' fit). However, because the SRMR represents an average of standardised residuals, it is important to inspect the residual correlation matrix to detect any unusual patterns or concerning values.

Alternative "goodness of fit" indices are typically less stringent. A widely used index is the root mean square error of approximation (RMSEA). The RMSEA is an "error of approximation" index because it determines the degree to which a model is a reasonable fit for the population. "Close" fit may be indicated by RMSEA values less than or equal to .05. A one-sided test calculates the probability value that RMSEA ≤ .05. Non-significant probability values (i.e., $p > .05$) may indicate acceptable model fit. Comparative fit indices, such as the Tucker-Lewis Index (TLI), evaluate the fit of a specified model in relation to a baseline "null" model, in which the covariances among all indicators are fixed to zero (Brown, 2015).

Factor loadings estimate the relationships between the factors and indicators. For example, a standardised factor loading estimate of 0.40 is interpreted as a standardised score increase in the factor associated with an 0.40 standardized score increase in the indicator. Squaring the factor loadings produces an estimate of the proportion of variance in the indicator accounted for by the latent factor (e.g., 0.40 = 16% variance explained). In factor analysis, the amount of variance in the indicator explained by the common factors is often referred to as the communality, whereas the remaining variance is unique (Kline, 2016).

I conducted the CFAs using the 'lavaan' package (Version 0.6-5; Rosseel, 2012).

Correlation Analyses

Because the SNAP-IV is a Likert scale, I calculated polychoric correlation coefficients to determine the relationships between items on the SNAP-IV subscales. SNAP-

IV and CBCL/6-18 subscale scores, however, are continuous. Therefore, I calculated Pearson correlation coefficients to determine the relationships between the SNAP-IV and CBCL/6-18 subscale scores. I also used the 'corcor' package to statistically compare Pearson correlation coefficients (Diedenhofen & Musch, 2015).

Internal Consistency

Internal consistency, as measured by coefficient alpha (also known as 'Cronbach's alpha') refers to the interrelatedness of items in a unidimensional scale (Cortina, 1993). Coefficient alpha is calculated using a matrix of Pearson correlation coefficients between all items in a particular scale or subscale (i.e., inter-item correlations). Ordinal coefficient alpha, using a polychoric correlation matrix, is a more reliable estimate of internal consistency for ordinal scales (Gadermann et al., 2012; Zumbo et al., 2007). I estimated the internal consistency reliability of the SNAP-IV subscales using ordinal coefficient alpha.

The standard for 'adequate' internal consistency depends on the context in which the scale is used, although applied researchers in the social sciences generally consider alpha coefficients of 0.70-0.79 to be 'adequate', 0.80-0.89 to be 'good', and greater than or equal to 0.90 to be 'excellent' (Kline, 2016).

Sample Size

I determined the minimum required sample sizes based on the statistical methods used for each analysis.

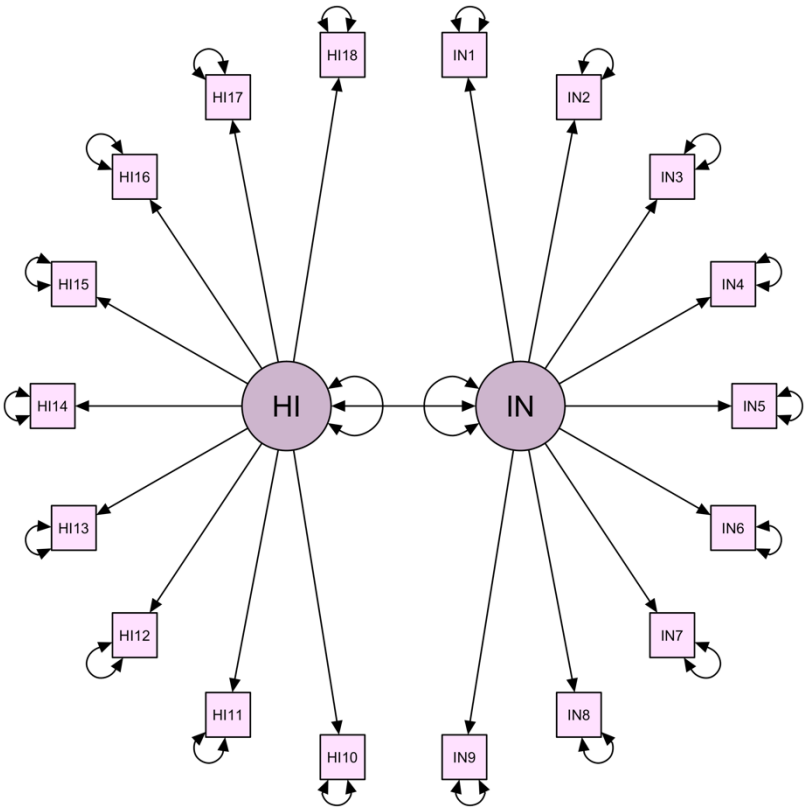
CFA

The minimally acceptable sample size required to conduct a CFA has been widely debated. Larger samples are always preferable for factor analysis; with larger samples, there is a reduced likelihood of error variance and improved ability to generalise the results. Model complexity, including the number of factors extracted and the number of indicators per factor, may influence the required sample size (Kyriazos, 2018). Sample sizes greater than 200 are often recommended, especially when models are complex, data are not normally distributed, and data are missing (Kline, 2016).

I conducted an a priori power analysis using the 'semPower' package in R (Version 1.1.0; Moshagen & Erdfelder, 2016). I determined the sample size required to detect a certain effect, in this case, a RMSEA less than or equal to 0.05, with a certain power, given a number of degrees of freedom. First, I calculated the degrees of freedom for the CFA of the SNAP-IV by subtracting the number of estimated parameters from the total number of variances and covariances (i.e., the data). Using the equation $\frac{p(p+1)}{2}$, where p is the number of items in a

tool, I determined that we had 171 ‘pieces’ of information. I then calculated the number of parameters that I wanted to estimate: Eight factor loadings per factor, corresponding to the nine items per subscale (one of the item loadings for each factor is set to 1.00), 20 variances (one per observed variable, and one for each latent factor), and one covariance between the two latent factors. I then subtracted the 37 parameters from 171 to obtain 134 degrees of freedom. Figure 1 displays the parameters to be estimated in the SNAP-IV CFA model. To obtain a RMSEA of 0.05 with 134 degrees of freedom and powers of 0.80, 0.70, and 0.60 respectively, sample sizes of 139, 119, and 102 are required. In the present study, we enrolled 109 participants.

Figure 1
Diagram Showing Parameters to be Estimated in CFA Model of the SNAP-IV



Note. CFA = confirmatory factor analysis. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition. HI = Hyperactivity-Impulsivity subscale. IN = Inattention subscale. Items are labelled as SubscaleItem (e.g., Item 1 = IN1). The straight arrows between latent factors and indicators represent factor loadings, the straight double-arrowed line between the two latent factors indicates covariance, and curved arrows represent variance. Diagram created with the semPlot package (Epskamp, 2019).

Correlation Analyses

A US-based study evaluating the psychometric properties of the SNAP-IV in children with ID found a large, positive correlation (r = .63) between total SNAP-IV scores and scores on the Inattention syndrome scale on the CBCL for Ages 4-18 (Miller et al., 2004). An a priori power analysis using G*Power software indicated that to achieve an effect size of 0.30 with 80% power for a one-tailed test ($\alpha = 0.05$), a minimum sample size of 64 is required (Faul et al., 2007). A small effect size of 0.30 was conservatively selected as smaller effect sizes are more common in this type of research (Rosnow & Rosenthal, 2003). In the present study, 79 of the 109 parents enrolled in the study completed the CBCL/6-18.

Ethical Considerations

Information regarding ethical approval, funding, and conflicts of interest are detailed in Chapter 4. Relevant ethical approval documents from the Human Research Ethics Council of the University of Cape Town, the Western Cape Department of Health, and the Institutional Review Board at Harvard T.H. Chan School of Public Health are presented in Appendices I-N.

Informed Consent

Consent processes commenced with a member of the data collection team introducing themselves, checking which language should be used (based on the participant's preference), making sure the setting was private and comfortable, and attending to any urgent questions or concerns from the parent. The team member then engaged the parent in a conversation about the research, using the information in the consent form to explain the study. The written consent forms were available in the participant's home language.

The consent form (Appendix O) included the following information:

i. A brief outline of the study purpose and detailed description of the study procedures, explained in simple language.
ii. Participants were free to withdraw from the study at any point, with no consequences.
iii. They would not receive feedback from their child's behavioural assessments.

All consent responses were captured on paper and electronically. Vocabulary and sentence structure in the consent forms were made as simple as possible to aid in easy understanding. Parents were encouraged to ask questions related to the information on the consent forms. No identifiable information from the paper consent forms were entered onto the electronic tablet. Two copies of the paper consent form were signed, one for the study records and one for the parent to keep.

Confidentiality and Anonymity

All participant data were coded with an alpha-numeric ID. In terms of data storage, all information loaded onto the REDCap system was encrypted on a password protected tablet and was deleted from the tablet once uploaded to the cloud storage system. To ensure anonymity, the alpha-numeric participant ID, which was not tied to any personal details (e.g., names and contact details from the paper consent form), was used to identify participants during data analysis. Assessments took place in a private space in one of the two tertiary hospitals to ensure confidentiality.

Risks and Benefits

Parents or guardians may have experienced distress or discomfort when answering questions about their child's condition(s) or experiences they have had with the child's development. They did not have to answer any questions that made them feel uncomfortable and were welcome to take a break at any time. If a participant expressed or exhibited any discomfort or distress, a psychologist on the research team was called to offer appropriate support. To prevent discomfort as much as possible, the research team took care to speak gently, to explain things clearly (and more than once if needed), and to create a safe space for parents without fear of being judged or criticised. There was a possibility that a child may have been embarrassed or uncomfortable hearing their parents answering questions about their behaviour (i.e., during administration of the SNAP-IV or the CBCL). To minimise this risk, these questionnaires were administered to the parent(s) in a separate room so that the child was not in the same room when these questions are being asked.

There were no direct benefits to participants for taking part in the study. If participants ask for feedback on the child's assessments, the researcher explained that feedback was not available (this was also stated clearly during the consent process).

References

Abubakar, A., Ssewanyana, D., de Vries, P. J., & Newton, C. R. (2016). Autism spectrum disorders in sub-Saharan Africa. *The Lancet Psychiatry*, *3*(9), 800–802. https://doi.org/10.1016/S2215-0366(16)30138-9

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. University of Vermont.

Amazue, L. O., Echezona, N. H., & Uzuegbu, C. N. (2015). Child temperament, gender, parental level of education and Attention Deficit-Hyperactivity Disorder among a Nigerian sample. *European Journal of Social Sciences*, *41*(1), 94–103.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Author.

Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, *45*(3), 601–613. https://doi.org/10.1017/S003329171400172X

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). The Guilford Press.

Bussing, R., Fernandez, M., Harwood, M., Hou, W., Garvan, C. W., Eyberg, S. M., & Swanson, J. M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: Psychometric properties and normative ratings from a school district sample. *Assessment*, *15*(3), 317–328. https://doi.org/10.1177/1073191107313888

Cluver, L. D., Lachman, J. M., Ward, C. L., Gardner, F., Peterson, T., Hutchings, J. M., Mikton, C., Meinck, F., Tsoanyane, S., Doubt, J., Boyes, M., & Redfern, A. A. (2017). Development of a parenting support program to prevent abuse of adolescents in South Africa: Findings from a pilot pre-post study. *Research on Social Work Practice*, *27*(7), 758–766. https://doi.org/10.1177/1049731516628647

Cortina, J. M. (1993). What Is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Costa, D. S., de Paula, J. J., Malloy-Diniz, L. F., Romano-Silva, M. A., & Miranda, D. M. (2019). Parent SNAP-IV rating of attention-deficit/hyperactivity disorder: Accuracy in a clinical sample of ADHD, validity, and reliability in a Brazilian sample. *Jornal de Pediatria*, *95*(6), 736–743. https://doi.org/10.1016/j.jpedp.2018.11.003

de Menil, V., Hoogenhout, M., Kipkemoi, P., Kamuya, D., Eastman, E., Galvin, A., Mwangasha, K., de Vries, J., Kariuki, S. M., Murugasen, S., Mwangi, P., Singh, I., Stein, D. J., Abubakar, A., Newton, C. R., Donald, K. A., & Robinson, E. (2019). The NeuroDev Study: Phenotypic and genetic characterization of neurodevelopmental disorders in Kenya and South Africa. *Neuron*, *101*(1), 15–19. https://doi.org/10.1016/j.neuron.2018.12.016

Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, *10*(4), Article e0121945. https://doi.org/10.1371/journal.pone.0121945

Epskamp, S. (2019). *semPlot: Path diagrams and visual analysis of various SEM packages' output* (R package version 1.1.2).

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*(4), 466–491. https://doi.org/10.1037/1082-989X.9.4.466

Fouché, N. (2017). *Putative genetic and environmental factors influencing Attention-Deficit/Hyperactivity Disorder (ADHD) in a South African sample* [University of the Free State]. https://scholar.ufs.ac.za/bitstream/handle/11660/8714/FoucheN.pdf?sequence=3&isAllowed=y

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, and Evaluation*, *17*(1), 3. https://doi.org/10.7275/N560-J767

Gau, S. S. F., Shang, C. Y., Liu, S. K., Lin, C. H., Swanson, J. M., Liu, Y. C., & Tu, C. L. (2008). Psychometric properties of the Chinese version of the Swanson, Nolan, and Pelham, version IV scale - Parent form. *International Journal of Methods in Psychiatric Research*, *17*(1), 35–44. https://doi.org/10.1002/mpr.237

Grañana, N., Richaudeau, A., Gorriti, C. R., O'Flaherty, M., Scotti, M. E., Sixto, L., Allegri, R., & Fejerman, N. (2011). Assessment of attention deficit hyperactivity: SNAP-IV scale adapted to Argentina. *Pan American Journal of Public Health*, *29*(5), 344–349. https://doi.org/10.1590/s1020-49892011000500007

Hall, C. L., Guo, B., Valentine, A. Z., Groom, M. J., Daley, D., Sayal, K., & Hollis, C. (2020). The validity of the SNAP-IV in children displaying ADHD symptoms. *Assessment*, *27*(6), 1258–1271. https://doi.org/10.1177/1073191119842255

Ivanova, M. Y., Dobrean, A., Dopfner, M., Erol, N., Fombonne, E., Fonseca, A. C., Frigerio, A., Grietens, H., Hannesdottir, H., Kanbayashi, Y., Lambert, M., Achenbach, T. M., Larsson, B., Leung, P., Liu, X., Minaei, A., Mulatu, M. S., Novik, T. S., Oh, K. J., … Chen, W. J. (2007). Testing the 8-syndrome structure of the Child Behavior Checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology*, *36*(3), 405–417. https://doi.org/10.1080/15374410701444363

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Kline, R. B. (2020). *Becoming a behavioral science researcher: A guide to producing research that matters* (2nd ed.). The Guilford Press.

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(8), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

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. https://doi.org/10.2989/17280583.2013.767804

Meyer, A., Eilertsen, D.-E., Sundet, J. M., Tshifularo, J., & Sagvolden, T. (2004). Cross-cultural similarities in ADHD-like behaviour amongst South African primary school children. *South African Journal of Psychology*, *34*(1), 122–138. https://doi.org/10.1177/008124630403400108

Miller, M. L., Fee, V. E., & Jones, C. J. (2004). Psychometric properties of ADHD rating scales among children with mental retardation II: Validity. *Research in Developmental Disabilities*, *25*(5), 477–492. https://doi.org/10.1016/j.ridd.2003.11.002

Moshagen, M., & Erdfelder, E. (2016). A new strategy for testing structural equation models. *Structural Equation Modeling*, *23*(1), 54–60. https://doi.org/10.1080/10705511.2014.950896

Myer, L., Stein, D. J., Grimsrud, A., Seedat, S., & Williams, D. R. (2008). Social determinants of psychological distress in a nationally-representative sample of South African adults. *Social Science & Medicine*, *66*(8), 1828–1840. https://doi.org/10.1016/j.socscimed.2008.01.025

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Revelle, W. (2020). *psych: Procedures for psychological, psychometric, and personality research* (R package version 2.0.12). Northwestern University. https://cran.r-project.org/web/packages/psych/index.html

Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, *57*(3), 221–237. https://doi.org/10.1037/h0087427

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Rust, J., & Golombok, S. (2009). *Modern psychometrics: The science of psychological assessment* (3rd ed.). Routledge.

Stein, D. J., Koen, N., Donald, K. A., Adnams, C. M., Koopowitz, S., Lund, C., Marais, A., Myers, B., Roos, A., Sorsdahl, K., Stern, M., Tomlinson, M., van der Westhuizen, C., Vythilingum, B., Myer, L., Barnett, W., Brittain, K., & Zar, H. J. (2015). Investigating the psychosocial determinants of child health in Africa: The Drakenstein Child Health Study. *Journal of Neuroscience Methods*, *252*, 27–35. https://doi.org/10.1016/j.jneumeth.2015.03.016

Swanson, J. M., Deutsch, C., Cantwell, D., Posner, M., Kennedy, J. L., Barr, C. L., Moyzis, R., Schuck, S., Flodman, P., Spence, M. A., & Wasdell, M. (2001). Genes and attention-deficit hyperactivity disorder. *Clinical Neuroscience Research*, *1*(3), 207–216. https://doi.org/10.1016/S1566-2772(01)00007-X

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., Clevenger, W., Wasdell, M., McCleary, R., Lakes, K., & Wigal, T. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN Rating Scales. *The International Journal of Educational and Psychological Assessment*, *10*(1), 51–70.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association.

Ullebø, A. K., Breivik, K., Gillberg, C., Lundervold, A. J., & Posserud, M.-B. (2012). The factor structure of ADHD in a general population of primary school children. *Journal of Child Psychology and Psychiatry*, *53*(9), 927–936. https://doi.org/10.1111/j.1469-7610.2012.02549.x

Visser, M. J., Hecker, H. E., & Jordaan, J. (2018). A comparative study of the psychological problems of HIV-infected and HIV-uninfected children in a South African sample. *AIDS Care*, *30*(5), 596–603. https://doi.org/10.1080/09540121.2017.1417530

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Zeegers, I., Rabie, H., Swanevelder, S., Edson, C., Cotton, M., & van Toorn, R. (2010). Attention Deficit Hyperactivity and Oppositional Defiance Disorder in HIV-infected South African children. *Journal of Tropical Pediatrics*, *56*(2), 97–102. https://doi.org/10.1093/tropej/fmp072

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. https://doi.org/10.22237/jmasm/1177992180

In the chapter that follows, I investigate the structural and construct validity of the Swanson, Nolan, and Pelham ADHD Rating Scale in a sample of South African children with neurodevelopmental disorders, including Autism Spectrum Disorder and Intellectual Disability. This study was submitted to a peer-reviewed journal and is currently under review. I conceptualised the study, collected, analysed and interpreted the data, and wrote the manuscript with input from my co-authors.

CHAPTER 4

Validity of the SNAP-IV for ADHD Assessment in South African Children with Neurodevelopmental Disorders

Michal R. Zieff[1], Michelle Hoogenhout[1,2], Emma Eastman[1], Björn U. Christ[1], Alice Galvin, Victoria de Menil[3,6], Amina Abubakar[5,6,7], Charles R. Newton[5,6,7], Elise Robinson[3,4], & Kirsten A. Donald[1]

[1]Department of Paediatrics and Child Health, University of Cape Town, Red Cross War Memorial Children's Hospital, Klipfontein Road, Rondebosch 7700, Cape Town, South Africa
[2]Galvanize, 44 Tehama Street, San Francisco, CA 94501, USA
[3]Stanley Centre for Psychiatric Research, Broad Institute of MIT and Harvard, 75 Ames Street, Cambridge, MA 02142, USA
[4]Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA
[5]Neurosciences Unit, Clinical Department, KEMRI-Wellcome Trust Collaborative Research Programme, PO Box 230-80108, Kilifi, Kenya
[6]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK
[7]Institute for Human Development, Aga Khan University, P.O. Box 30270-00100 Nairobi, Kenya

Conflicts of Interest and Source of Funding

The authors declare no conflicts of interest. The NeuroDev study is supported by the Stanley Centre for Psychiatric Research at the Broad Institute and the National Institute of Mental Health (1U01MH119689-01). The NeuroDev Study also received funding through a grant from the Simons Foundation/SFARI (599648, ER and KAD). In addition, MRZ received fellowship funding from the Harry Crossley Foundation.

Acknowledgements

We extend our gratitude to the families who participated in the study. We also acknowledge the important work of the NeuroDev South Africa data collection team.

Abstract

Objective The Swanson, Nolan, and Pelham ADHD Rating Scale (SNAP-IV) is frequently used to evaluate symptoms of Attention-Deficit/Hyperactivity Disorder (ADHD). Only a few studies have investigated the validity of the SNAP-IV in children with other neurodevelopmental disorders. This study investigated the psychometric properties of the SNAP-IV in a sample of South African children with neurodevelopmental disorders. *Method* Parents of 109 children completed the SNAP-IV. A subset of parents (n = 79) also completed the Child Behaviour Checklist (CBCL/6-18). We conducted a confirmatory factor analysis to inspect the two-factor structure of the SNAP-IV (Inattention + Hyperactivity-Impulsivity). We also calculated ordinal coefficient alpha (α) to estimate internal consistency. Finally, we correlated scores of SNAP-IV and CBCL/6-18 subscales to estimate concurrent, convergent (ADHD + externalising behaviour), and discriminant (ADHD + internalising behaviour) validity respectively. *Results* The two-factor model performed acceptably. The model fit improved after removing three problematic items, two of which were dependent on the child's level of spoken language. The revised SNAP-IV subscales had acceptable internal consistencies (α = 0.85-0.86). Correlation coefficients between the SNAP-IV and ADHD-related CBCL/6-18 subscales were significant (r = 0.53-0.62, $p < 0.001$). Correlations between ADHD and externalising behaviours (r = 0.45, $p < 0.001$) and internalising behaviours (r = 0.38, $p < 0.001$) respectively were not significantly different (z = 0.97, p = 0.165). *Conclusion* The findings tentatively support the use of the SNAP-IV in this group of children. However, there are limitations to its performance in this population likely related to the presence of NDDs.

Keywords: SNAP-IV, Attention-Deficit/Hyperactivity Disorder, neurodevelopmental disorders, psychometric, behaviour rating scale

Attention-Deficit/Hyperactivity Disorder (ADHD) frequently occurs as a comorbidity in children with other neurodevelopmental disorders (NDDs), including Autism Spectrum Disorder (ASD), Intellectual Disability (ID), and Specific Learning Disorder (SLD). Clinically significant ADHD symptoms may be present in approximately 30% to 78% of children with ASD, 23% of children with ID, and 40% of children with SLDs (Bandla et al., 2017; Gadow et al., 2005; Lee & Ousley, 2006; Leyfer et al., 2006; Neece et al., 2011; Rao & Landa, 2014; Sinzig et al., 2009; Smith & Adams, 2006; Stevens et al., 2016).

Notwithstanding the frequent co-occurrence of ADHD and other NDDs, less is known about the nature of ADHD symptoms in children with other NDDs than about ADHD symptoms in children who are otherwise typically developing (Reilly & Holland, 2011). One pertinent issue is that of 'diagnostic overshadowing', whereby symptoms of other behavioural or mental health disorders tend to be overlooked in children with ID, or are attributed to the ID diagnosis (Reiss & Szyszko, 1983). For example, clinicians may attribute ADHD-related behaviours (e.g., difficulty sustaining attention or organizing tasks) to deficits in intellectual and adaptive functioning, or may decide that ADHD-related behaviours are developmentally appropriate or consistent with a child's developmental age (American Psychiatric Association, 2013). Similar concerns have been raised about the manifestation of ADHD-related behaviours in children with ASD (Frazier et al., 2001; Lee & Ousley, 2006). Diagnostic overshadowing makes it difficult to determine whether ADHD-related behaviours manifest in a similar or comparable manner in children with other NDDs compared to those considered otherwise to be typically developing. This, in turn, makes it difficult to assess whether our approach to measuring ADHD-related behaviours in children with other NDDs is valid.

Behavioural rating scales are often used in clinical settings to identify problem behaviours associated with childhood disorders, including ADHD. Data derived from ADHD rating scales may play an important supportive role in clinical assessment. For example, clinicians may rely on rating scales to quantify the severity of impairments, to monitor changes in behaviour over time or monitor impact of interventions (Young et al., 2020). It is therefore important that the information derived from these measures is valid and reliable. Children with other NDDs are frequently excluded from normative and validation samples of behavioural measures designed to evaluate symptoms of ADHD. This is likely a consequence of our limited understanding of ADHD in this population and the difficulty of assigning behavioural symptoms to one or other condition (Antshel et al., 2006). The result is a lack of

evidence demonstrating that ADHD rating scales measure the same underlying constructs (i.e., have construct validity) with children who have other NDDs.

There is limited evidence suggesting that self-report instruments designed to detect ADHD symptoms may not be valid for children with NDDs. One study evaluating the ADHD Rating Scale-Fourth Edition in a sample of children with ASD, without accompanying intellectual disability, found that the scale failed to adequately distinguish between inattention and hyperactivity-impulsivity, the two constructs thought to underpin ADHD in this group (Yerys et al., 2017). A confirmatory factor analysis (CFA) found three items intended to measure inattention, including "Does not listen when spoken to directly" were also associated with the hyperactivity-impulsivity factor in these children. These results suggest that the presence of comorbid ASD may influence the ratings of target (ADHD) behaviours. However, in terms of the overall phenotype described, higher ADHD scores were associated with higher levels of externalising behaviours (e.g., aggression) relative to internalising behaviours (e.g., depression), resembling findings from typically developing samples (Reiersen & Todorov, 2013).

One frequently used ADHD rating scale is the Swanson, Nolan, and Pelham ADHD Rating Scale (4th edition; SNAP-IV; Swanson et al., 2012) Like many other ADHD rating scales, the SNAP-IV items are based on the Diagnostic and Statistical Manual of Mental Disorders (5th edition; DSM-5) criteria for ADHD (American Psychiatric Association, 2013). In research contexts, the SNAP-IV is typically used with school-based, non-clinical samples (Bussing et al., 2008). However, a few promising studies support the use of the SNAP-IV with children who have ID. Two related studies found that the SNAP-IV subscales had strong psychometric properties in a sample of children with ID. The SNAP-IV had excellent reliability and concurrent validity, demonstrated by large, positive correlations with scores on other ADHD rating scales (Miller, Fee, & Jones, 2004; Miller, Fee, & Netterville, 2004).

In clinical contexts, the SNAP-IV ADHD rating scale is sometimes used to aid in the assessment of ADHD in child patients with ASD, ID and other NDDs. The SNAP-IV is used clinically in this way in the Western Cape province of South Africa, where, similar to other resource-limited settings, specialist neurodevelopmental services accessible to the majority of the population are frequently over-burdened. In addition, the clinical severity of neurodevelopmental populations presenting to tertiary hospitals in Africa is typically high, resulting in a large number of non-verbal children (Bakare & Munir, 2011). Rating scales are especially useful for assessment purposes in low-resourced clinical settings as they are quick and inexpensive to administer. However, until there is sufficient psychometric evidence to

support use in these contexts, results should be interpreted with caution. To the best of our knowledge, there are no published studies to date that investigate the psychometric properties of the SNAP-IV in a sample of children with ASD, including those who are non-verbal and who may have associated ID.

The aim of this study was to evaluate the validity of scores derived from the SNAP-IV in a sample of young South African children with NDDs. We hypothesized that (i) the SNAP-IV items would measure two distinguishable constructs, namely, inattention and hyperactivity-impulsivity, (ii) the SNAP-IV subscale scores would correlate positively with subscale scores of another tool measuring ADHD-related behaviours, and (iii) the SNAP-IV would have strong, positive correlations with externalising behaviours relative to internalising behaviours.

Methods

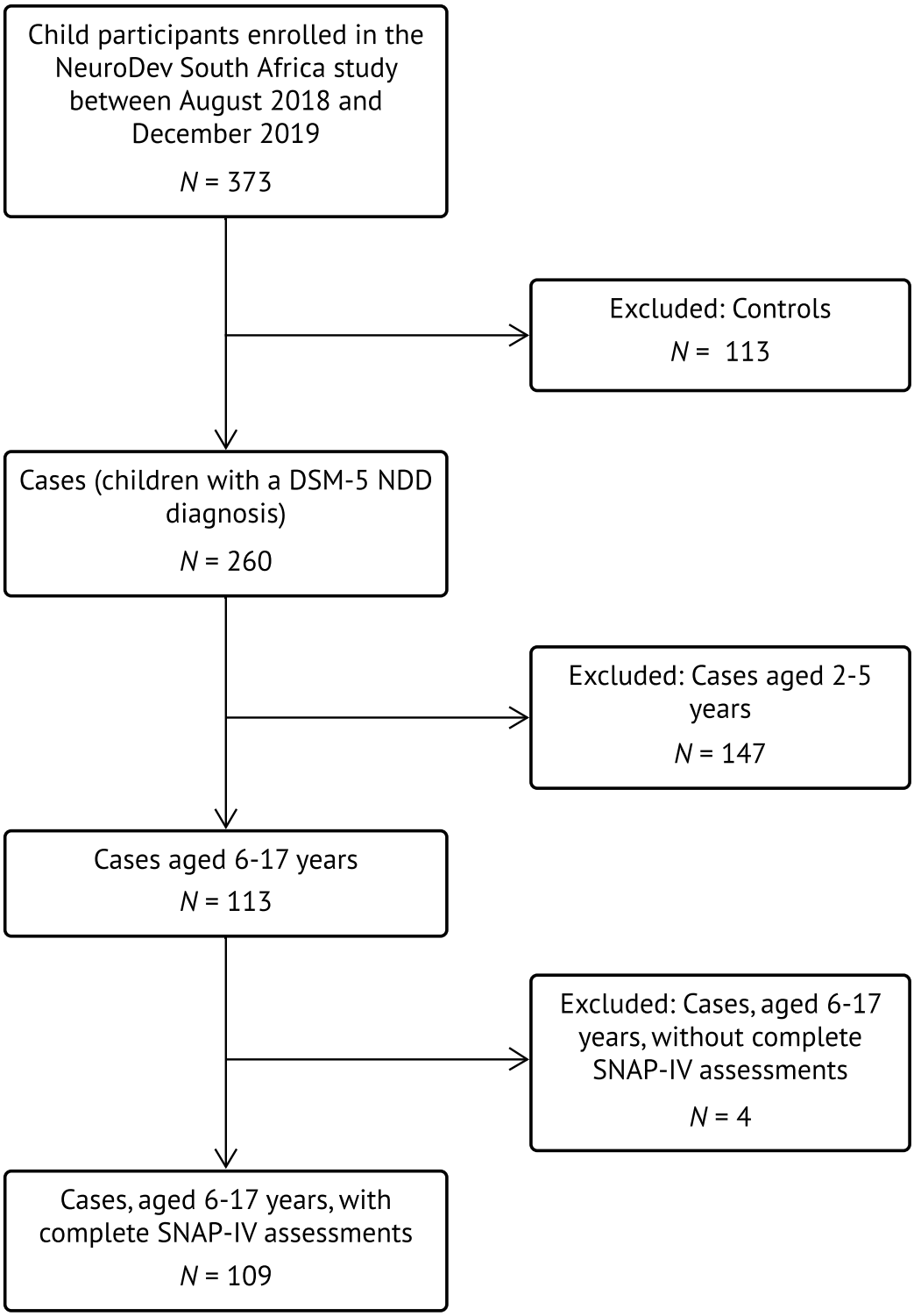
Participants

This study was embedded within a larger project titled "NeuroDEV South Africa", which administers the SNAP-IV to parents of case children (children with NDDs) aged 6-17 years (de Menil et al., 2019). Figure 1 outlines the sample selection process for the current study within the context of the larger NeuroDEV South Africa study.

We invited parents of children attending outpatient developmental, genetic, speech, and neurology clinics at two tertiary hospitals in Cape Town to participate in the NeuroDEV South Africa study. Children were eligible for inclusion as cases if they had a confirmed diagnosis of one or more of the following DSM-5 NDDs: ADHD, ASD, SLD, ID or Communication Disorder (CD). Children with a primary neuromotor disability (e.g., cerebral palsy) were not eligible to participate. We analysed data from all case children aged 6 years and older who were enrolled in the NeuroDEV South Africa study between August 2018 and December 2019, with the exception of four participants who did not complete the SNAP-IV assessments.

Figure 1

Sample Selection for the Current Study Within the NeuroDEV South Africa Study



Note. DSM-5 = Diagnostic and Statistical Manual of Mental Disorders, 5th edition. NDD = neurodevelopmental disorder. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition. Figure created with Lucidchart.com.

Measures

Demographics Questionnaire and Asset Index

Parents completed a demographic questionnaire, which included questions about the child and parents' home languages. The Asset Index is a socio-demographic questionnaire adapted from previous South African studies, including questions regarding parental educational attainment and household income (Myer et al., 2008; Stein et al., 2015).

Swanson, Nolan and Pelham ADHD Rating Scale (4th edition; SNAP-IV)

The SNAP-IV ADHD Rating Scale is an 18-item self-report questionnaire designed to measure symptoms of ADHD (Swanson et al., 2012). The items are consistent with the DSM-5 ADHD diagnostic criteria and are designed to distinguish between different symptom presentations of ADHD, namely, inattentive, hyperactive-impulsive, and combined (both inattentive and hyperactive/impulsive). The subscales are named accordingly; 'Inattention' (IN, 9 items) and 'Hyperactivity-Impulsivity' (HI, 9 items). The items are presented as DSM-5 symptoms (e.g., "Often is forgetful in daily activities"). For each item, respondents select one of four response options (0 "Not at all", 1 "Just a little", 2 "Quite a bit", or 3 "Very much") that best describes the child's behaviour over the past year. Subscale and total scores are calculated as an average score across relevant items. We included a "not applicable" (N/A) response option to capture questions relating to speech in non-verbal children.

We obtained Afrikaans and Xhosa translations of the SNAP-IV from another South African research group that had previously used the SNAP-IV (Zeegers et al., 2010). Two Afrikaans and three Xhosa-speaking members of the NeuroDEV South Africa research team independently reviewed the Afrikaans and Xhosa translations respectively to confirm that the translations were tapping into the intended constructs. The Afrikaans review team consisted of a senior research assistant with a neuropsychology background (BUC) and a paediatric research nurse. The Xhosa review team comprised a senior research assistant with a neuropsychology background, a postgraduate student with a psychology background (MRZ), and a paediatric research nurse. Each team consolidated their suggestions for revisions (minor revisions to the wording of items) to create the final translated versions.

To estimate the prevalence of clinically significant ADHD symptoms in this sample, we set cut-off scores that were aligned with the DSM-5 diagnostic criteria for ADHD. The DSM-5 requires an individual to exhibit at least 6 symptoms in at least one domain (inattention or hyperactivity-impulsivity) to qualify for a diagnosis of ADHD. Pelham and colleagues (1992) recommend defining the presence of a symptom by a 3-point score on the SNAP-IV items, as it produces prevalence rates similar to those reported in the general

population. Therefore, we estimated the prevalence of ADHD symptomology in the sample by calculating the number of participants who scored a '3' on at least six items in at least one domain.

Child Behaviour Checklist for Ages 6-18 (CBCL/6-18)

The CBCL/6-18 is a self-report questionnaire designed to assess specific problematic behaviours in school-age children, as reported by caregivers (Achenbach & Rescorla, 2001). It is considered a "gold standard" tool for assessing behavioural problems in children and has been validated in over 30 countries (Ivanova et al., 2007). The items listed in the CBCL/6-18 are designed to align with the DSM-5 diagnostic criteria for a number of behavioural disorders, including ADHD. This study was concerned with four subscales: The Attention Problems syndrome scale (10 items), the ADHD DSM-oriented scale (7 items), the Aggressive Behaviour syndrome scale (18 items), and the Withdrawn/Depressed syndrome scale (8 items). We used the latter two subscales as measures of externalising and internalising behaviours respectively. We obtained licenses to administer the ASEBA Afrikaans and Xhosa translations of the CBCL/6-18. For each item, parents rated the frequency/severity of their child's behaviour in the past six months using a three-point scale; 0 ('not true'), 1 ('somewhat or sometimes true'), or 2 ('very or often true'). Subscale scores were calculated by summing the scores of all relevant items.

Procedure

We obtained informed consent from parents of child participants. If developmentally appropriate, we obtained assent from child participants over the age of 12 years. A member of the data collection team (comprising medical doctors, research nurses, and psychologists) verbally administered the SNAP-IV subscales to all parents (N = 109) as well as the CBCL/6-18 subscales to a subset of parents (n = 79), in their preferred language. Although we did not record the language of administration for specific tools, anecdotal reports indicate that the majority of participants elected to complete the questionnaires in English.

The University of Cape Town's Human Research Ethics Committee approved this research (367/2019) as part of the NeuroDEV South Africa study (810/2016). The NeuroDEV study also received ethics approval from the Harvard T. H. Chan School of Public Health Institutional Review Board (17-1260) and the Western Cape Department of Health, South Africa.

Statistical analyses

We performed all statistical analyses in RStudio (Version 1.3.1093) for R (Version 4.0.2; R Core Team, 2019). We used the 'psych' package to conduct exploratory data

analysis (Revelle, 2020). To estimate the concurrent validity, convergent validity, and divergent validity of the SNAP-IV, we computed Pearson's correlation coefficients between SNAP-IV subscale scores and CBCL/6-18 subscale scores, and the 'corcor' package to statistically compare correlation coefficients (Diedenhofen & Musch, 2015). We estimated the internal consistency reliability of the SNAP-IV subscales using ordinal coefficient alpha (Zumbo et al., 2007). To test the expected two-factor structure of the SNAP-IV, we conducted a CFA using the 'lavaan' package (Version 0.6-5; Rosseel, 2012). CFA examines relationships between observed variables (indicators) and latent variables (factors). CFA is used to confirm an a priori specification (i.e., hypothesis) about the underlying structure of a tool (Kline, 2016).

We used global and local fit statistics to determine the usefulness of the model (Brown, 2015). The chi square statistic (χ^2) is a test of exact fit between the model and the data. The Tucker-Lewis Index (TLI) conveys information about the "goodness of fit" between the model and the data. A higher value indicates better fit, and should ideally exceed 0.95 (i.e., the specified model should improve the fit by 95% relative to no model). The Root Mean Square Error of Approximation (RMSEA) is a measure of approximate fit, where 0 indicates a perfect fit and 0.05 indicates close fit. The Standardised Root Mean Square Residual (SRMR) represents the average difference between the observed and predicted correlations matrices (i.e., the average residual correlation). Like the RMSEA, SRMR values close to 0 suggest good fit. Modification Indices (MIs) approximate the amount by which chi-square value will decrease (i.e., model fit will improve) if an unspecified parameter were to be estimated. Parameters with MIs greater than 3.84 indicate a statistically significant decrease in the chi-square statistic, or improved model fit. Each MI is associated with a standardized expected parameter change (SEPC), the magnitude and direction of which approximates how much the parameter is expected to change if it were to be estimated. Standardised coefficients (also known as 'factor loadings') are estimates of direct effects between latent and observed variables. The squared factor loadings indicate the proportion of variance in each observed variable that is explained by the latent factor.

Results

The sample included 109 child participants aged 6-15 years (M = 7.54, SD = 1.97). Table 1 presents the sociodemographic characteristics of the sample, as well as information related to DSM-5 diagnoses, level of spoken language, and highest level of education. Most children came from low- and middle-income families that spoke one of the three main

languages spoken in the Western Cape of South Africa. The predominance of male participants (> 70%) is typical of study samples comprising children with NDDs (Springer et al., 2013).

Fifty-three child participants (49%) had more than one NDD diagnosis. The most frequent diagnoses were ID and ASD. ASD and ID were comorbid in 32 participants (29% of the sample). In this sample, CD was primarily comorbid with ID, with only two children having CD as a primary diagnosis. Eighteen children (17%) were diagnosed with ADHD, which in all but one participant was comorbid with at least one other NDD, including ASD (*n* = 4), ID (*n* = 8), ASD and ID (*n* = 3), CD (*n* = 1), or SLD (*n* = 1). Seventy-five children (69%) had delayed (i.e., non-fluent) speech. Forty-three (39%) children were not enrolled in a formal schooling system.

Table 1

Sociodemographic and Diagnostic Information (N = 109)

Variable	Frequency (%)	
Sex		
Male	86	(77.90)
Female	23	(21.10)
Home language		
English	70	(64.22)
Afrikaans	10	(9.17)
Xhosa	24	(22.02)
Other	5	(4.59)
Asset Index		
Monthly household income		
< $66	2	(1.89)
$66 –$333	42	(39.62)
$333 – $666	32	(30.19)
$666 – $1000	15	(14.15)
> $1000	12	(11.32)
Unknown	3	(2.83)
Highest level of maternal education		
No education	1	(0.94)
Primary school	4	(3.77)
High school without completion	36	(33.96)
Completed high school	28	(26.42)
Tertiary education (partial or complete)	36	(33.96)
Unknown	1	(0.94)
Overall DSM-5 diagnoses (including comorbid diagnoses)*		
Attention-Deficit/Hyperactivity Disorder	18	(16.51)
Autism Spectrum Disorders	68	(62.39)
Intellectual Disability	69	(63.30)
Communication Disorders	8	(7.34)
Specific Learning Disorders	1	(0.92)
Language level		
No spoken language	26	(23.85)
Single words only	17	(15.60)
Phrases	32	(29.36)
Fluent	34	(31.19)
Highest level of education		
Never attended	23	(21.10)
Crèche	20	(18.35)
Non-academic/alternative curriculum	14	(12.84)
Special needs school	19	(17.43)
Pre-primary school	13	(11.93)
Primary school (Grades 1-6)	19	(17.43)
Other	1	(0.91)

Note. DSM-5 = Diagnostic and Statistical Manual of Mental Disorders, 5th edition. 'Other' home languages included Shona (n = 3), Swahili (n = 1), and Lingala (n = 1). All participants who spoke the aforementioned 'other' languages also spoke English at home. There were 106 responses to the Asset Index as there were three sets of siblings in the sample. At the time of writing, $1.00 ≈ ZAR15.00. Information about the child's diagnosis, language level, and education were extracted from the child's medical records by a medical officer.

*Percentages do not add up to 100 as children may have more than one diagnosis.

Table 2 presents a summary of the sample's SNAP-IV scores, as well as item response frequencies. Figures S1 and S2 in the supplement display the distribution of subscale scores. Forty children (37%) exhibited clinically significant symptoms of ADHD. Fourteen children (13%) met the cut-off for a predominantly inattentive presentation of ADHD, 16 (15%) for a predominantly hyperactive/impulsive presentation, and 10 (9%) for a combined presentation. Of the 18 participants with a confirmed ADHD diagnosis, 3 (17%) met the criterion for a predominantly inattentive presentation, 3 (17%) for a predominantly hyperactive/impulsive presentation, and 4 (22%) for a combined presentation. There were no significant differences in the numbers of children who met the criteria for different presentations of ADHD, χ^2 (2, N = 40) = 1.37, p = 0.497. On average, respondents strongly endorsed the SNAP-IV items, with percentages of "quite a bit" or "very much" responses ranging from 20-77% (M = 59.48, SD = 15.56) and percentages of "very much" responses ranging from 9-61% (M = 37.31, SD = 12.87). Response patterns did not differ substantially by primary diagnosis (see Figure S3 in the supplement). Seven items had small proportions of N/A responses, while Items 15 ("Talks excessively") and 16 ("Blurts out answers") had 31 (28%) and 33 (30%) N/A responses respectively.

Table 2

SNAP-IV Item Frequencies, Means, and Standard Deviations ($N = 109$)

SNAP-IV Item		Response frequencies (%)					Percentage responses		*M* (*SD*)
		"0"	"1"	"2"	"3"	"N/A"	% "2" or "3"	% "3"	
Inattention									
1	Fails to give close attention to details	10 (9.17)	25 (22.94)	38 (34.86)	35 (32.11)	1 (0.92)	66.97	32.11	1.91 (0.96)
2	Difficulty sustaining attention	13 (11.93)	18 (16.51)	31 (28.44)	47 (43.12)	0 (0.00)	71.56	43.12	2.03 (1.04)
3	Does not listen when spoken to directly	5 (4.59)	30 (27.52)	44 (40.37)	29 (26.61)	1 (0.92)	66.97	26.61	1.90 (0.85)
4	Fails to finish tasks	7 (6.42)	23 (21.10)	35 (32.11)	42 (38.53)	2 (1.83)	70.64	38.53	2.05 (0.94)
5	Difficulty organizing tasks and activities	24 (22.02)	18 (16.51)	21 (19.27)	44 (40.37)	2 (1.83)	59.63	40.37	1.79 (1.20)
6	Avoids tasks requiring sustained effort	13 (11.93)	16 (14.68)	21 (19.27)	57 (52.29)	2 (0.00)	71.56	52.29	2.14 (1.08)
7	Loses things necessary for tasks	41 (37.61)	13 (11.93)	22 (20.18)	33 (30.28)	0 (0.00)	50.46	30.28	1.43 (1.27)
8	Distracted by extraneous stimuli	19 (17.43)	13 (11.93)	25 (22.94)	52 (47.71)	0 (0.00)	70.64	47.71	2.01 (1.14)
9	Forgetful in daily activities	35 (32.11)	21 (19.27)	25 (22.94)	27 (24.77)	1 (0.92)	47.71	24.77	1.41 (1.18)
Hyperactivity-Impulsivity									
10	Fidgets or squirms	19 (17.43)	17 (15.60)	22 (20.18)	51 (46.79)	0 (0.00)	66.97	46.79	1.96 (1.15)
11	Leaves seat	18 (16.51)	19 (17.43)	24 (22.02)	47 (43.12)	1 (0.92)	65.14	43.12	1.93 (1.13)
12	Runs about or climbs excessively	30 (27.52)	12 (11.01)	22 (20.18)	45 (41.28)	0 (0.00)	61.47	41.28	1.75 (1.26)
13	Difficulty playing quietly	52 (47.71)	21 (19.27)	14 (12.84)	22 (20.18)	0 (0.00)	33.03	20.18	1.06 (1.19)
14	"On the go" as if "driven by a motor"	24 (22.02)	16 (14.68)	24 (22.02)	45 (41.28)	0 (0.00)	63.30	41.28	1.83 (1.19)
15	Talks excessively	17 (15.60)	20 (18.35)	14 (12.84)	27 (24.77)	31 (28.44)	37.61	24.77	1.65 (1.17)
16	Blurts out answers	36 (33.03)	18 (16.51)	12 (11.01)	10 (9.17)	33 (30.28)	20.18	9.17	0.95 (1.08)
17	Difficulty awaiting turn	15 (13.76)	10 (9.17)	17 (15.60)	67 (61.47)	0 (0.00)	77.06	61.47	2.25 (1.10)
18	Interrupts or intrudes on others	16 (14.68)	17 (15.60)	24 (22.02)	52 (47.71)	0 (0.00)	69.72	47.71	2.03 (1.11)

Note. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition. Response categories "0", "1", "2", "3" are labelled as "Not at all", "A little bit", "Quite a bit", and "Very much" respectively, while "N/A" indicates "Not applicable" responses. Some wordings of the SNAP-IV items are abbreviated.

Average polychoric correlation coefficients (ρ, see Table 3) between items on the Inattention subscale (M = 0.38, SD = 0.15) and between items on the Hyperactivity-Impulsivity subscale (M = 0.36, SD = 0.19) respectively were only slightly larger than cross-subscale correlations (M = 0.30, SD = 0.16), tentatively suggesting a weak distinction between the two constructs. We noted one negative correlation coefficient between two Hyperactivity-Impulsivity items, Items 13 ("Difficulty playing quietly") and 16 ("Blurts out answers"; ρ = -0.15, n = 76), and a few near-zero correlation coefficients between three pairs of Hyperactivity-Impulsivity items, 11 ("Leaves seat") and 15 ("Talks excessively"; ρ = -0.01, n = 78), 13 ("Difficulty playing quietly") and 15 ("Talks excessively"; ρ = 0.02, n = 78), 11 ("Leaves seat") and 16 ("Blurts out answers"; ρ = 0.07, n = 76), and between one pair of Inattention Items, 5 ("Difficulty organizing") and 8 ("Easily distracted"; ρ = 0.05, n = 107).

Table 3

Polychoric Correlation Matrix and Item-Total Correlations for the SNAP-IV (N = 109)

SNAP-IV Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	ITC
Inattention																		
Item 1																		0.60
Item 2	0.60																	0.61
Item 3	0.42	0.39																0.51
Item 4	0.51	0.54	0.47															0.58
Item 5	0.53	0.50	0.38	0.51														0.53
Item 6	0.69	0.55	0.35	0.60	0.51													0.62
Item 7	0.45	0.36	0.29	0.28	0.33	0.48												0.42
Item 8	0.11	0.44	0.30	0.26	0.05	0.24	0.10											0.26
Item 9	0.23	0.35	0.44	0.40	0.44	0.36	0.47	0.33										0.50
Hyperactivity-Impulsivity																		
Item 10	0.31	0.40	0.33	0.35	0.28	0.34	0.22	0.40	0.44									0.44
Item 11	0.55	0.45	0.38	0.43	0.29	0.56	0.45	0.22	0.27	0.37								0.48
Item 12	0.44	0.37	0.45	0.38	0.27	0.37	0.40	0.26	0.41	0.27	0.71							0.59
Item 13	0.47	0.27	0.35	0.32	0.36	0.52	0.39	0.14	0.43	0.49	0.54	0.59						0.43
Item 14	0.23	0.19	0.34	0.15	0.18	0.30	0.30	0.21	0.41	0.54	0.32	0.49	0.46					0.50
Item 15	-0.05	0.05	0.18	0.13	0.07	-0.15	-0.04	0.16	0.40	0.24	-0.01	0.19	0.02	0.47				0.33
Item 16	-0.08	0.11	0.28	0.16	0.15	-0.18	0.06	0.41	0.29	0.15	0.07	0.37	-0.15	0.44	0.57			0.34
Item 17	0.39	0.48	0.48	0.39	0.41	0.32	0.20	0.16	0.32	0.39	0.51	0.57	0.48	0.27	0.22	0.21		0.47
Item 18	0.18	0.25	0.54	0.23	0.27	0.41	0.56	0.42	0.50	0.42	0.35	0.39	0.35	0.49	0.36	0.48	0.36	0.48

Note. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition. ITC = Item-total correlation. Correlation coefficients within the black borders are cross-subscale correlation coefficients (i.e., between items on the Inattention and Hyperactivity-Impulsivity subscales). Some correlation coefficients are based on $n < 109$ due to non-applicable responses (see Table 2). ITCs were calculated using total subscale scores (Inattention OR Hyperactivity-Impulsivity), not total SNAP-IV scores. For example, ITC for Item 1 represents the Pearson correlation coefficient between Item 1 scores and Inattention subscale scores (minus Item 1).

We conducted a CFA specifying a model (Model 1) with two latent factors, Inattention and Hyperactivity-Impulsivity, each with nine indicators (SNAP-IV Items 1-9 and Items 10-18 respectively), as well as an additional path estimating the covariance between the two latent factors. Table 4 presents details of the model fit and Table 5, the standardized coefficients (factor loadings). The approximate model fit indices (TLI, RMSEA, and SRMR) were approaching acceptable levels. There was a significant correlation of 0.791 ($p < 0.001$) between the two latent factors. In general, the factors explained relatively high proportions of variance within their respective items. Exceptions were Items 15 ("Talks excessively") and 16 ("Blurts out answers"), of which Hyperactivity-Impulsivity explained only 11% and 13% variance respectively, and Item 8 ("Easily distracted"), of which Inattention explained only 17% variance.

Table 4

Model Fit Statistics for Confirmatory Factor Analysis Models (N = 109)

Model	χ^2 (df)	TLI	RMSEA (90% CI)	SRMR
1[a]	249.82 (134) ***	0.868	0.089 (0.072-0.107) ***	0.123
2[b]	196.05 (118) ***	0.909	0.078 (0.058-0.097) *	0.109
3[c]	158.17 (103) ***	0.934	0.070 (0.048-0.092)	0.092
4[d]	135.17 (89) **	0.942	0.069 (0.044-0.092)	0.087

Note. TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation, 90% CI = 90% Confidence intervals, SRMR = Standardized Root Mean Square Residual. Both chi-square (χ^2) and RMSEA significance tests are accept-support tests. All models measure two latent variables, 'Inattention' ('IN') and 'Hyperactivity-Impulsivity' ('HI'). Covariance between the two latent variables is also specified. Estimation method is robust weighted least squares with polychoric correlations, which is recommended for ordinal data (Li, 2016).

Estimate is significant at *$p < 0.05$, ** $p < 0.01$, or ***$p < 0.001$.

[a]Model 1: All 18 items – IN (9) and HI (9)

[b]Model 2: Item 15 removed – IN (9) and HI (8)

[c]Model 3: Items 15 and 16 removed – IN (9) and HI (7)

[d]Model 4: Items 15, 16, and 8 removed – IN (8) and HI (7)

Table 5

Communalities (R^2) and Standardized Coefficients (Factor Loadings) of SNAP-IV Items in Two CFA Models (N = 109)

SNAP-IV Item	Communalities / Factor loadings					
	Model 1 (Initial)			Model 4 (Final)		
	R^2	IN	HI	R^2	IN	HI
Inattention						
1. Fails to give close attention	0.555	0.745		0.581	0.762	
2. Difficulty sustaining attention	0.501	0.708		0.488	0.699	
3. Does not listen when spoken to	0.428	0.654		0.415	0.644	
4. Fails to finish tasks	0.464	0.681		0.463	0.680	
5. Difficulty organizing tasks	0.391	0.626		0.405	0.637	
6. Avoids effortful tasks	0.604	0.777		0.631	0.794	
7. Loses things	0.352	0.593		0.371	0.609	
8. Distracted by extraneous stimuli	0.173	0.416		-	-	
9. Forgetful in daily activities	0.428	0.654		0.399	0.632	
Hyperactivity-Impulsivity						
10. Fidgets or squirms	0.391		0.625	0.370		0.609
11. Leaves their seat	0.587		0.766	0.608		0.780
12. Runs about excessively	0.600		0.775	0.594		0.770
13. Difficulty playing quietly	0.485		0.696	0.524		0.724
14. Often "on the go"	0.370		0.608	0.334		0.578
15. Talks excessively	0.111		0.333	-		-
16. Blurts out answers	0.131		0.362	-		-
17. Difficulty awaiting turn	0.455		0.675	0.458		0.677
18. Interrupts or intrudes on others	0.434		0.659	0.390		0.624

Note. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition, CFA = confirmatory factor analysis, IN = Inattention subscale, HI = Hyperactivity-Impulsivity subscale. Factor loadings are standardized.

An inspection of the modification indices and a residual correlation matrix revealed many correlated residuals (r > absolute value of 0.1; see Figure S4, a residual correlation matrix for Model 1, and Table S1, both in the supplement). For example, Item 16 ("Blurts out answers") and Item 13 ("Difficulty playing quietly") had a large negative residual correlation (MI = 29.18, r = -0.40), suggesting that the model parameters are overestimating the relationship between these two items. However, the large positive residual correlation between Items 15 ("Talks excessively") and 16 ("Blurts out answers"; MI = 10.06, r = 0.45) suggests that the specified model is not fully accounting for the covariance between these two items. MIs suggested possible "cross-loadings" of Items 3 ("Does not listen"; MI = 6.00, SEPC = 0.457) and 9 ("Forgetful in daily activities", MI = 5.02, SEPC = 0.436) with the Hyperactivity-Impulsivity factor. We compared Model 1 to a nested one-factor model (Model 1a) to determine if variance in the SNAP-IV indicators are better explained by a single latent factor, "ADHD". The factor loadings of the three poorer items identified in Model 1 (Items 15, 16, and 8) were still low, and the model fit was poorer than that of the original two-factor model, χ^2 (135, N = 109) = 277.62, p < .001, TLI = 0.849, RMSEA = 0.099.

Taken together, these findings indicate that Items 15 ("Talks excessively") and 16 ("Blurts out answers") are poor items in this sample. Hence, we sequentially removed Items 15 (Model 2) and 16 (Model 3) from the initial model. The model fit improved substantially with each removal (see Table 4). We also removed Item 8 ("Easily distracted") as the Inattention factor did not seem to be explaining much variance in the item. The final model (Model 4) specified two related latent factors, Inattention and Hyperactivity-Impulsivity, with 8 and 7 indicators respectively (see Figure S5 in the supplement). Factor loadings corresponding to the remaining items did not differ substantially from the initial model (see Table 5). There were a few relatively large negative residual correlations between items on the same subscale (e.g., Items 1/9, r = -0.25; Items 10/12, r = -0.20; see Figure S6, a residual correlation matrix for Model 4, in the supplement). MIs suggesting possible cross-loadings of Items 3 ("Does not listen"; MI = 5.83, SEPC = 0.481) and 9 ("Forgetful"; MI = 4.04, SEPC = 0.418, see Table S1 in the supplement) with Hyperactivity-Impulsivity remained significant. The model fit, however, was excellent and the RMSEA was not significantly different from 0.05 (p = 0.096). Items 8, 15, and 16 were excluded from all further analyses.

Given evidence to support two correlated unidimensional factors, we calculated subscale means as well as internal consistency reliability statistics for the new Inattention and Hyperactivity-Impulsivity subscales respectively. The means were identical for Inattention (M = 1.83, SD = 0.71), Hyperactivity-Impulsivity (M = 1.83, SD = 0.77) and Combined

ADHD (M = 1.83, SD = 0.67) subscales respectively. The distributions of subscale scores were approximately normal (skewness = -0.36, -0.27, and -0.38 for IN, HI, and C subscales respectively. Ordinal alpha coefficients were good for Inattention (α = 0.86, 95% CI = 0.83-0.90), Hyperactivity-Impulsivity (α = 0.85, 95% CI = 0.81-0.89), and Combined ADHD (α = 0.86, 95% CI = 0.83-0.90) respectively.

Table S2 in the supplement presents a summary of CBCL/6-18 scores for the subset of the sample that completed the CBCL/6-18 (n = 79). Inattention (SNAP-IV) and Attention Problems (CBCL/6-18) scores were moderately correlated (r = 0.53, 95% CI 0.35-0.67, p < 0.001). Likewise, there was a moderate, positive correlation between scores on the SNAP-IV Combined ADHD subscale and the CBCL/6-18 ADHD DSM-oriented subscale (r = 0.57, 95% CI 0.40-0.70, p < 0.001). Combined ADHD (SNAP-IV) scores were also significantly correlated with scores on the CBCL/6-18 Aggression Problems subscale (r = 0.48, 95% CI = 0.29-0.64, p < 0.001), and the CBCL/6-18 Withdrawn/Depressed scores (r = 0.38, 95% CI = 0.17-0.55, p < 0.001). The difference between these two correlation coefficients was not statistically significant (z = 0.97, p = 0.165).

Discussion

The primary aim of the study was to investigate the construct validity of the SNAP-IV in a sample of children with NDDs. A CFA tentatively supported the claim that the scale measures two factors, Inattention and Hyperactivity-Impulsivity, barring a few problematic items which did not seem to measure either factor clearly.

The two "verbal" items, Items 15 ("Talks excessively") and 16 ("Blurts out answers"), had notably small factor loadings. In this sample, excessive talking was likely an indicator of a child's verbal abilities, rather than a measure of hyperactive/impulsive behaviour. From a statistical perspective, the large number of N/A responses for these items, likely due to the substantial proportion of the sample with limited expressive language, may also have contributed to the items' weak factor loadings. We were not able to use multiple imputation techniques to counteract the effects of missing data points, as the data were not missing at random. The large negative residual correlation between Items 15 and 16 may further account for the items' poor performance in the model. The model did not sufficiently account for the observed variance shared between these two items. It is possible that similar response patterns for these items may have contributed to the unexpectedly large amount of shared variance (i.e., a caregiver of a non-verbal child who responded to Item 15 with N/A likely responded to Item 16 in the same way).

Our understanding of Item 8's ("Distracted by extraneous stimuli") relatively poor performance in the models is less clear. Item 8 had a smaller average inter-item correlation with other items on the Inattention subscale, and a smaller item-total correlation, suggesting that the item was not measuring the same construct as other items in the subscale. One possible explanation is that the wording of the item may have been too technical, which in turn may have influenced the interpretation of the item. Although the data collection team encouraged participants to ask for clarification if needed, it is possible that some caregivers were reluctant to ask for clarification or responded acquiescently. A simplification of the item wording may be recommended to reduce the likelihood of misunderstandings and inaccurate responses.

Moderate, positive relationships between the SNAP-IV and CBCL/6-18 ADHD-related subscales supported the concurrent validity of the SNAP. It is worth noting that the SNAP-IV and CBCL/6-18 are similar in terms of item content and response format. Hence it is possible that the correlation coefficients may have been inflated by common method variance. The SNAP-IV did not demonstrate convergent and discriminant validity when correlating ADHD-related behaviours with internalising and externalising behaviours respectively. Combined ADHD (SNAP-IV) scores had equally weak associations with withdrawn/depressed and aggressive behaviour, as measured by the CBCL/6-18. These results should be interpreted with caution, as it is likely that sample-related factors may have affected responses to the CBCL/6-18. The size of the sub-sample that completed the CBCL/6-18 was too small to conduct a robust psychometric evaluation of the CBCL/6-18. However, it is possible that items such as "Teases a lot", "Threatens people" (Aggressive Behaviour), "Refuses to talk" and "Secretive" (Withdrawn/Depressed) were not applicable to children who were non-verbal. Perhaps the pre-school version of the CBCL, the CBCL/1.5-5 may have been a more appropriate measure of internalising and externalising behaviours, given the average developmental level of the children in this sample.

The DSM-5 requires that symptoms of ADHD be inconsistent with an individual's developmental level. The SNAP-IV was designed for use with typically developing (i.e., non-clinical) school-aged children without serious comorbid conditions (Swanson et al., 2012). The average child in the current sample was chronologically young, had some degree of language delay, and was not attending a mainstream school. Some of the SNAP-IV items, especially those indicating inattention (e.g., "Avoids tasks requiring sustained mental effort"), may not have been relevant to children not yet enrolled in a schooling system where such behaviours are typically expected. In other words, the behaviours indicated by the items

(e.g., sustaining attention, modulating verbal activity) may not have been appropriate to expect of a child, taking into account their developmental age and the presence of specific cognitive deficits. The SNAP-IV may therefore be less useful as a measure of ADHD for children with moderate to severe developmental delay. Notwithstanding the promising results of the CFA (especially given the significant differences between the 'target' sample and the current sample), the clinical context in which the tool is being administered has important implications for the interpretation of items and responses.

Limitations

This study had three major limitations. Most of the children with ASD had some degree of cognitive or language delay. Although the sample was representative of children attending neurodevelopmental clinics in sub-Saharan Africa (Bakare & Munir, 2011; Springer et al., 2013), these results cannot be generalized to NDD populations with different proportions of severe and non-verbal cases. A larger sample of children with NDDs, stratified by diagnosis, chronological age, and degree of language delay would have allowed a more thorough evaluation of the validity of the SNAP-IV in the NDD population. Another important limitation was the small sample size. CFA techniques often require large sample sizes to ensure precision and generalizability of the results (Kline, 2016). Chi-square tests, fit indices, parameter estimates, modification indices, and standardized residuals are all sensitive to sample size (Kyriazos, 2018). Larger samples (>200) are often recommended, especially when models are complex, data are not normally distributed, and when data are missing (Kline, 2016). A sample size of at least 200 would likely produce a more stable solution. Finally, including a variety of behavioural measures with different measurement methods would have likely provided more accurate and reliable estimates of convergent and discriminant validity of the SNAP-IV.

Summary and Conclusion

Clinicians working in low-resource settings in the Western Cape often administer the SNAP-IV to evaluate ADHD symptoms in children with NDDs. Data derived from the SNAP-IV have important implications for referral and the provision of interventions for comorbid ADHD. However, little is known about its validity in the NDD population, especially in children with moderate to severe ID and language delay as part of their phenotype. The primary aim of this study was to evaluate the validity of the SNAP-IV in a sample of children with NDDs using CFA techniques. The results of the CFA tentatively supported the validity of the SNAP-IV in this sample. However, there were limitations to its performance related to the clinical characteristics of the sample. The presence of

developmental delay and specific cognitive deficits associated with one or more NDDs likely influenced ratings of target ADHD behaviours. Importantly, two SNAP-IV items may be unsuitable measures of hyperactivity-impulsivity in children without functional speech. Other items, especially those measuring inattention, may be less relevant to children not yet enrolled in a formal schooling system. Additional studies are needed to determine whether existing DSM-based ADHD rating scales are able to capture inattentive and hyperactive/impulsive behaviours in the NDD population with sufficient precision.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. University of Vermont.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.

Antshel, K. M., Phillips, M. H., Gordon, M., Barkley, R., & Faraone, S. V. (2006). Is ADHD a valid disorder in children with intellectual delays? *Clinical Psychology Review*, *26*(5), 555–572. https://doi.org/10.1016/j.cpr.2006.03.002

Bakare, M. O., & Munir, K. M. (2011). Excess of non-verbal cases of autism spectrum disorders presenting to orthodox clinical practice in Africa – a trend possibly resulting from late diagnosis and intervention. *South African Journal of Psychiatry*, *17*(4), 118–120. https://doi.org/10.4102/sajpsychiatry.v17i4.295

Bandla, S., Mandadi, G. D., & Bhogaraju, A. (2017). Specific learning disabilities and psychiatric comorbidities in school children in South India. *Indian Journal of Psychological Medicine*, *39*(1), 76–82. https://doi.org/10.4103/0253-7176.198950

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). The Guilford Press.

Bussing, R., Fernandez, M., Harwood, M., Hou, W., Garvan, C. W., Eyberg, S. M., & Swanson, J. M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: Psychometric properties and normative ratings from a school district sample. *Assessment*, *15*(3), 317–328. https://doi.org/10.1177/1073191107313888

de Menil, V., Hoogenhout, M., Kipkemoi, P., Kamuya, D., Eastman, E., Galvin, A., Mwangasha, K., de Vries, J., Kariuki, S. M., Murugasen, S., Mwangi, P., Singh, I., Stein, D. J., Abubakar, A., Newton, C. R., Donald, K. A., & Robinson, E. (2019). The NeuroDev Study: Phenotypic and genetic characterization of neurodevelopmental disorders in Kenya and South Africa. *Neuron*, *101*(1), 15–19. https://doi.org/10.1016/j.neuron.2018.12.016

Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, *10*(4), Article e0121945. https://doi.org/10.1371/journal.pone.0121945

Frazier, J. A., Biederman, J., Bellordre, C. A., Garfield, S. B., Geller, D. A., Coffey, B. J., & Faraone, S. V. (2001). Should the diagnosis of Attention-Deficit/ Hyperactivity disorder be considered in children with Pervasive Developmental Disorder? *Journal of Attention Disorders*, *4*(4), 203–211. https://doi.org/10.1177/108705470100400402

Gadow, K. D., Devincent, C. J., Pomeroy, J., & Azizian, A. (2005). Comparison of DSM-IV symptoms in elementary school-age children with PDD versus clinic and community samples. *Autism*, *9*(4), 392–415. https://doi.org/10.1177/1362361305056079

Ivanova, M. Y., Dobrean, A., Dopfner, M., Erol, N., Fombonne, E., Fonseca, A. C., Frigerio, A., Grietens, H., Hannesdottir, H., Kanbayashi, Y., Lambert, M., Achenbach, T. M., Larsson, B., Leung, P., Liu, X., Minaei, A., Mulatu, M. S., Novik, T. S., Oh, K. J., … Chen, W. J. (2007). Testing the 8-syndrome structure of the Child Behavior Checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology*, *36*(3), 405–417. https://doi.org/10.1080/15374410701444363

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(8), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Lee, D. O., & Ousley, O. Y. (2006). Attention-deficit hyperactivity disorder symptoms in a clinic sample of children and adolescents with pervasive developmental disorders. *Journal of Child and Adolescent Psychopharmacology*, *16*(6), 737–746. https://doi.org/10.1089/cap.2006.16.737

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., Tager-Flusberg, H., & Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, *36*(7), 849–861. https://doi.org/10.1007/s10803-006-0123-0

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Miller, M. L., Fee, V. E., & Jones, C. J. (2004). Psychometric properties of ADHD rating scales among children with mental retardation II: Validity. *Research in Developmental Disabilities*, *25*(5), 477–492. https://doi.org/10.1016/j.ridd.2003.11.002

Miller, M. L., Fee, V. E., & Netterville, A. K. (2004). Psychometric properties of ADHD rating scales among children with mental retardation I: Reliability. *Research in Developmental Disabilities*, *25*(5), 459–476. https://doi.org/10.1016/j.ridd.2003.11.003

Myer, L., Stein, D. J., Grimsrud, A., Seedat, S., & Williams, D. R. (2008). Social determinants of psychological distress in a nationally-representative sample of South African adults. *Social Science & Medicine*, *66*(8), 1828–1840. https://doi.org/10.1016/j.socscimed.2008.01.025

Neece, C. L., Baker, B. L., Blacher, J., & Crnic, K. A. (2011). Attention-deficit/hyperactivity disorder among children with and without intellectual disability: An examination across time. *Journal of Intellectual Disability Research*, *55*(7), 623–635. https://doi.org/10.1111/j.1365-2788.2011.01416.x

Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*(2), 210–218. https://doi.org/10.1097/00004583-199203000-00006

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rao, P. A., & Landa, R. J. (2014). Association between severity of behavioral phenotype and comorbid attention deficit hyperactivity disorder symptoms in children with autism spectrum disorders. *Autism*, *18*(3), 272–280. https://doi.org/10.1177/1362361312470494

Reiersen, A. M., & Todorov, A. A. (2013). Exploration of ADHD subtype definitions and co-occurring psychopathology in a Missouri population-based large sibship sample. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *1*(1), 3–13. https://doi.org/10.21307/sjcapp-2013-002

Reilly, C., & Holland, N. (2011). Symptoms of Attention Deficit Hyperactivity Disorder in children and adults with Intellectual Disability: A review. *Journal of Applied Research in Intellectual Disabilities*, *24*(4), 291–309. https://doi.org/10.1111/j.1468-3148.2010.00607.x

Reiss, S., & Szyszko, J. (1983). Diagnostic overshadowing and professional experience with mentally retarded persons. *American Journal of Mental Deficiency*, *87*(4), 396–402. https://psycnet.apa.org/record/1983-10672-001

Revelle, W. (2020). *psych: Procedures for psychological, psychometric, and personality research* (R package version 2.0.12). Northwestern University. https://cran.r-project.org/web/packages/psych/index.html

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Sinzig, J., Walter, D., & Doepfner, M. (2009). Attention Deficit/Hyperactivity Disorder in children and adolescents with Autism Spectrum Disorder: Symptom or syndrome? *Journal of Attention Disorders*, *13*(2), 117–126. https://doi.org/10.1177/1087054708326261

Smith, T., & Adams, G. (2006). The effect of comorbid AD/HD and learning disabilities on parent-reported behavioral and academic outcomes of children. *Learning Disability Quarterly*, *29*(2), 101–112. https://doi.org/10.2307/30035538

Springer, P. E., Van Toorn, R., Laughton, B., & Kidd, M. (2013). Characteristics of children with pervasive developmental disorders attending a developmental clinic in the Western Cape Province, South Africa. *South African Journal of Child Health*, *7*(3), 95–99. https://doi.org/10.7196/sajch.530

Stein, D. J., Koen, N., Donald, K. A., Adnams, C. M., Koopowitz, S., Lund, C., Marais, A., Myers, B., Roos, A., Sorsdahl, K., Stern, M., Tomlinson, M., van der Westhuizen, C., Vythilingum, B., Myer, L., Barnett, W., Brittain, K., & Zar, H. J. (2015). Investigating the psychosocial determinants of child health in Africa: The Drakenstein Child Health Study. *Journal of Neuroscience Methods*, *252*, 27–35. https://doi.org/10.1016/j.jneumeth.2015.03.016

Stevens, T., Peng, L., & Barnard-Brak, L. (2016). The comorbidity of ADHD in children diagnosed with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *31*, 11–18. https://doi.org/10.1016/j.rasd.2016.07.003

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., Clevenger, W., Wasdell, M., McCleary, R., Lakes, K., & Wigal, T. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN Rating Scales. *The International Journal of Educational and Psychological Assessment*, *10*(1), 51–70.

Yerys, B. E., Nissley-Tsiopinis, J., de Marchena, A., Watkins, M. W., Antezana, L., Power, T. J., & Schultz, R. T. (2017). Evaluation of the ADHD Rating Scale in youth with Autism. *Journal of Autism and Developmental Disorders*, *47*(1), 90–100. https://doi.org/10.1007/s10803-016-2933-z

Young, S., Hollingdale, J., Absoud, M., Bolton, P., Branney, P., Colley, W., Craze, E., Dave, M., Deeley, Q., Farrag, E., Gudjonsson, G., Hill, P., Liang, H. L., Murphy, C., Mackintosh, P., Murin, M., O'Regan, F., Ougrin, D., Rios, P., … Woodhouse, E. (2020). Guidance for identification and treatment of individuals with attention deficit/hyperactivity disorder and autism spectrum disorder based upon expert consensus. *BMC Medicine*, *18*, Article 146. https://doi.org/10.1186/s12916-020-01585-y

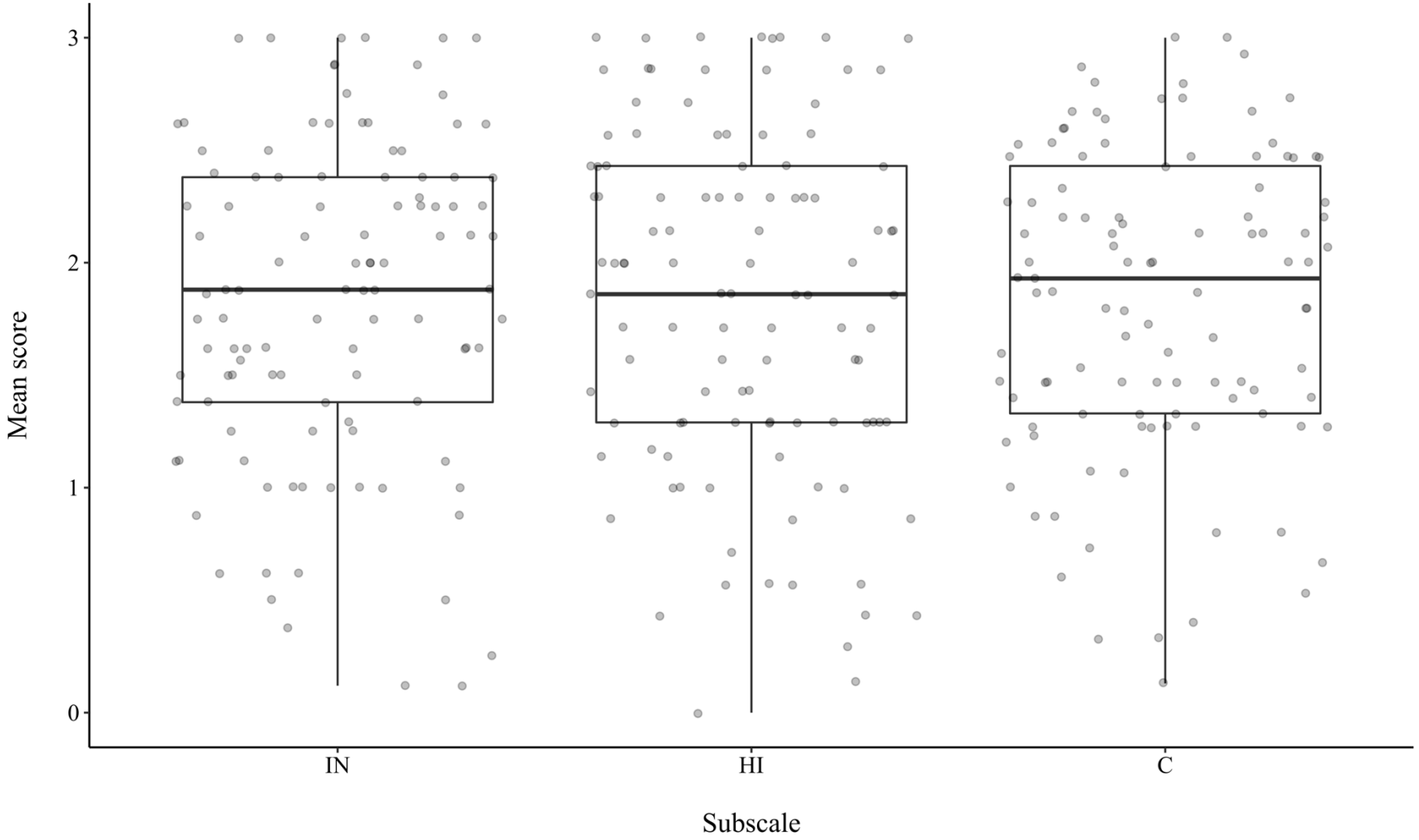
Zeegers, I., Rabie, H., Swanevelder, S., Edson, C., Cotton, M., & van Toorn, R. (2010). Attention Deficit Hyperactivity and Oppositional Defiance Disorder in HIV-infected South African children. *Journal of Tropical Pediatrics*, *56*(2), 97–102. https://doi.org/10.1093/tropej/fmp072

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. https://doi.org/10.22237/jmasm/1177992180

Supplementary Materials

Figure S1

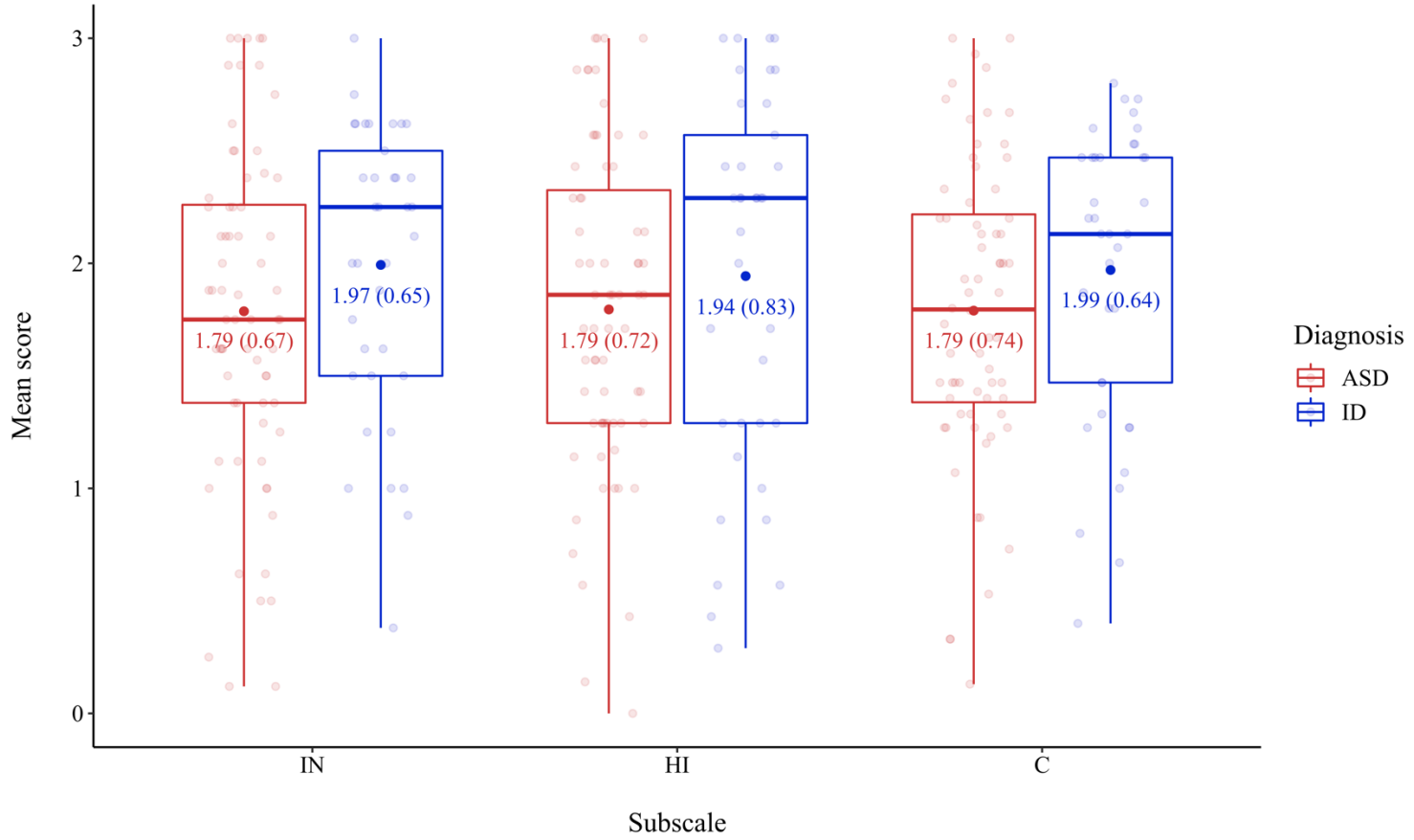
Distribution of Mean Scores on SNAP-IV Subscales for Total Sample (N = 109)



Note. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition. IN = Inattention subscale, HI = Hyperactivity-Impulsivity subscale, C = Combined scale.

Figure S2

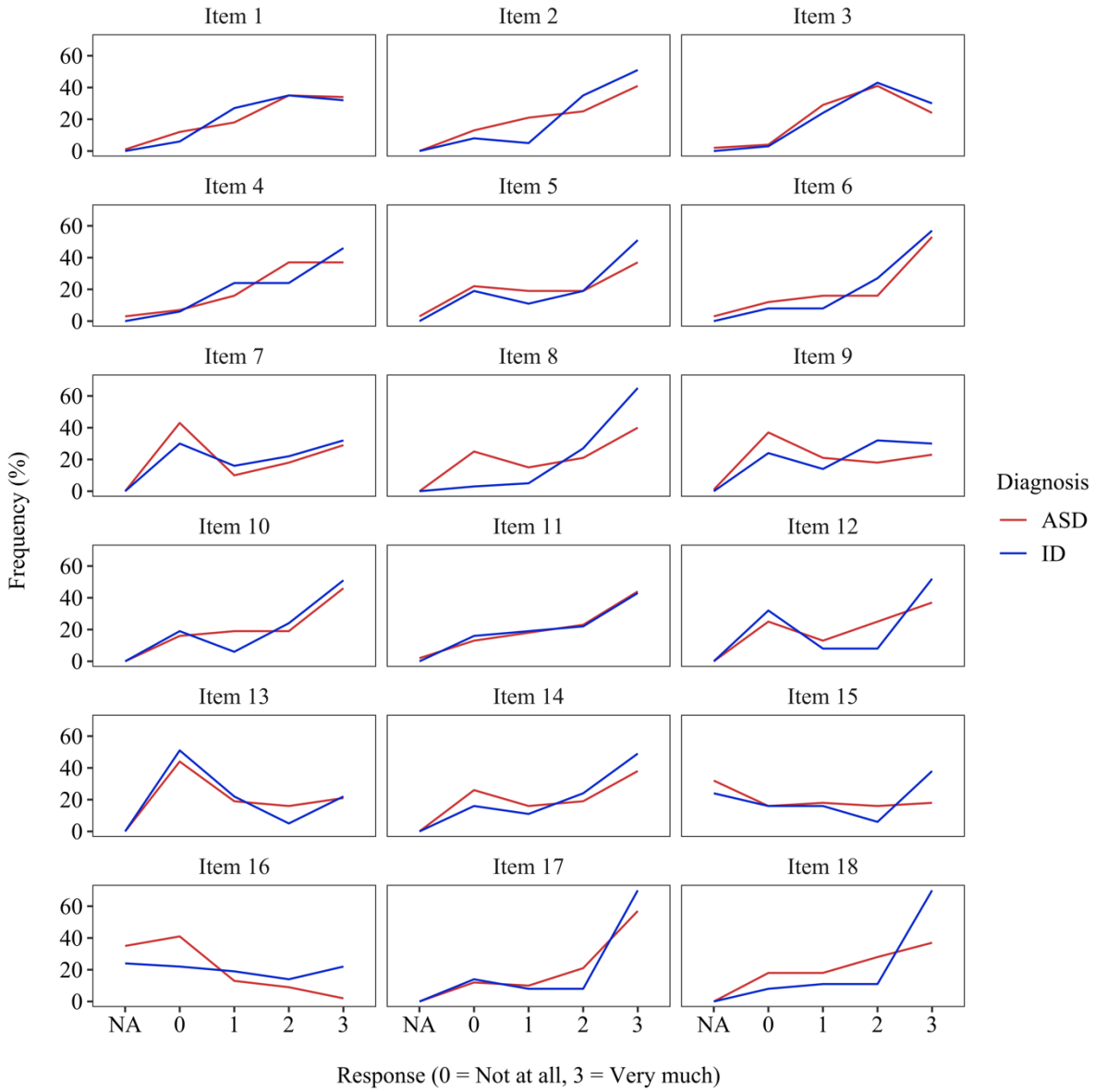
Distribution of Mean Scores of the SNAP-IV Subscales by Diagnosis (N = 105)



Note. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition. ASD = Autism Spectrum Disorder. ID = Intellectual Disability. For the purposes of this plot, the data represent only those participants with a primary diagnosis of ASD (n = 68) or ID (n = 37). Participants with comorbid ASD and ID are grouped under the ASD category. Data from the remaining participants with a primary diagnosis of Communication Disorder (n = 2), Attention Deficit/Hyperactivity Disorder (n = 1), and Specific Learning Disorder (n = 1), are excluded. The filled dots displayed in the horizontal centre of the boxes represent mean scores. Mean values are reported below the dots with standard deviations in parentheses.

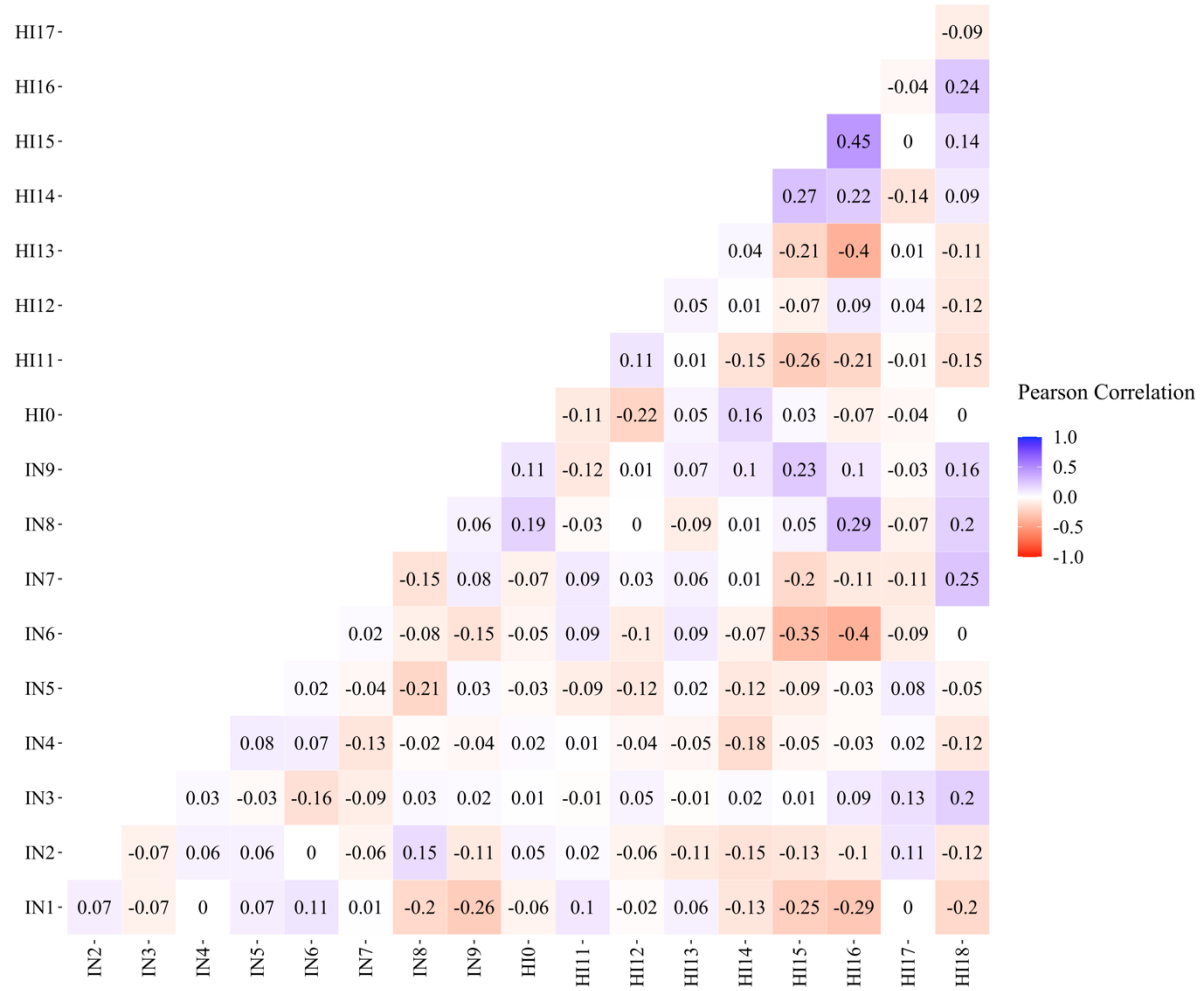
Figure S3

Response Frequencies (%) for SNAP-IV Items by Diagnostic Group (N = 105)



Note. SNAP-IV = Swanson, Nolan, and Pelham ADHD Rating Scale, 4th edition. ASD = Autism Spectrum Disorder. ID = Intellectual Disability. For the purposes of this plot, the data represent only those participants with a primary diagnosis of ASD (n = 68) or ID (n = 37). Participants with comorbid ASD and ID are grouped under the ASD category. Data from the remaining participants with a primary diagnosis of Communication Disorder (n = 2), Attention Deficit/Hyperactivity Disorder (n = 1), and Specific Learning Disorder (n = 1) are excluded.

Figure S4
Residual Correlation Matrix for CFA Model 1 (N = 109)



Note. CFA = confirmatory factor analysis. IN = Inattention subscale, HI = Hyperactivity-Impulsivity subscale. Items are labelled as SubscaleItem (e.g., Item 1 = IN1).

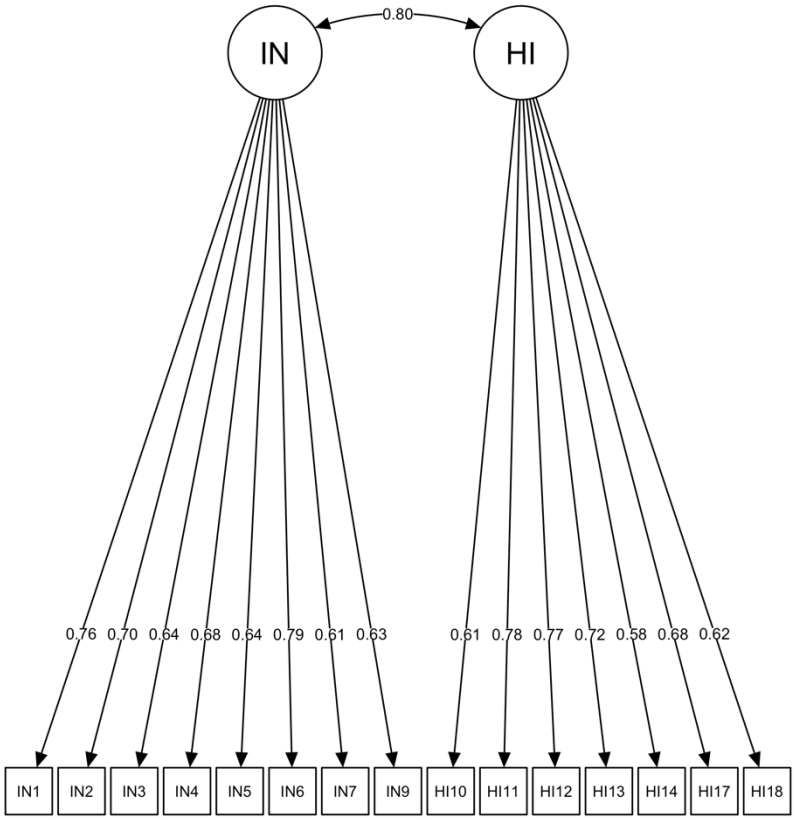
Table S1

Modification Indices and Expected Parameter Changes for CFA Models 1 and 4 (N = 109)

	Factor loadings				Residual correlations			
Model	Factor	Item	MI	SEPC	Item	Item	MI	SEPC
1								
	HI	IN-3	6.00	0.457	HI-15	HI-16	29.18	0.560
	HI	IN-9	5.02	0.436	HI-13	HI-16	10.06	-0.661
					IN-6	HI-16	9.64	-0.733
					IN-7	HI-18	8.01	0.455
					HI-14	HI-15	7.57	0.406
4								
	HI	IN-3	5.83	0.481	IN-7	HI-18	8.33	0.454
	HI	IN-9	4.04	0.418	IN-3	HI-18	8.13	0.425
					IN-1	IN-9	6.51	-0.585
					HI-10	HI-14	5.35	0.341
					HI-11	HI-12	4.83	0.466

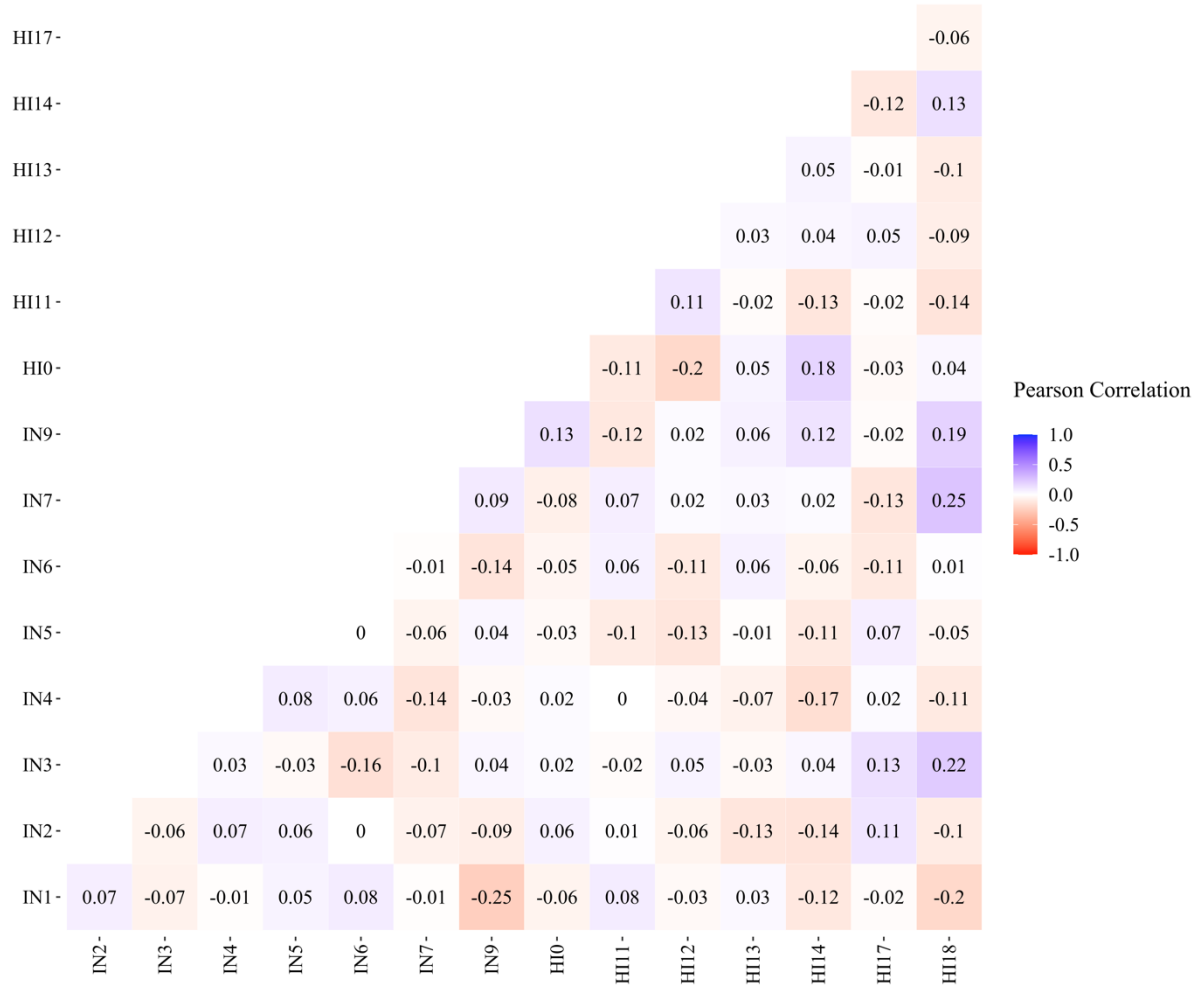
Note. CFA = confirmatory factor analysis, MI = modification index, SEPC = standardised expected parameter change, IN = Inattention subscale, HI = Hyperactivity-Impulsivity subscale. Items are labelled as SubscaleItem (e.g., Item 1 = IN1). Only the five largest significant modification indices are displayed. For the 'factor loadings', only MIs with positive SEPCs are reported.

Figure S5
Diagram of CFA Model 4 (N = 109)



Note. CFA = confirmatory factor analysis. IN = Inattention subscale, HI = Hyperactivity-Impulsivity subscale. Items are labelled as SubscaleItem (e.g., Item 1 = IN1). Path labels are standardised coefficients (factor loadings).

Figure S6
Residual correlation matrix for CFA Model 4 (N = 109)



Note. CFA = confirmatory factor analysis. IN = Inattention subscale, HI = Hyperactivity-Impulsivity subscale. Items are labelled as SubscaleItem (e.g., Item 1 = IN1).

Table S2

CBCL/6-18 item frequencies, means, and standard deviations (*N* = 79)

CBCL/6-18 Subscale/Item		Response frequencies (%)			*M (SD)*
		"0"	"1"	"2"	
Attention Problems Syndrome Scale					8.57 (3.06)
1	Acts too young for his/her age	13 (16.46)	20 (25.32)	46 (58.23)	1.42 (0.76)
4	Fails to finish things he/she starts	11 (13.92)	30 (37.97)	38 (48.10)	1.34 (0.71)
8	Can't concentrate	5 (6.33)	30 (37.97)	44 (55.70)	1.49 (0.62)
10	Can't sit still	12 (15.19)	25 (31.65)	42 (53.16)	1.38 (0.74)
13	Confused or seems to be in a fog	39 (49.37)	31 (39.24)	9 (11.39)	0.62 (0.69)
17	Daydreams	30 (37.97)	31 (39.24)	18 (22.78)	0.85 (0.77)
41	Impulsive or acts without thinking	20 (25.32)	29 (36.71)	30 (37.97)	1.13 (0.79)
61	Poor school work	52 (65.82)	14 (17.72)	13 (16.46)	0.51 (0.77)
78	Inattentive or easily distracted	6 (7.59)	19 (24.05)	54 (68.35)	1.61 (0.63)
80	Stares blankly	36 (45.57)	32 (40.51)	11 (13.92)	0.68 (0.71)
ADHD DSM-Oriented Scale					11.03 (3.74)
4	Fails to finish things he/she starts	11 (13.92)	30 (37.97)	38 (48.10)	1.34 (0.71)
8	Can't concentrate	5 (6.33)	30 (37.97)	44 (55.70)	1.49 (0.62)
10	Can't sit still	12 (15.19)	25 (31.65)	42 (53.16)	1.38 (0.74)
41	Impulsive or acts without thinking	20 (25.32)	29 (36.71)	30 (37.97)	1.13 (0.79)
78	Inattentive or easily distracted	6 (7.59)	19 (24.05)	54 (68.35)	1.61 (0.63)
93	Talks too much	44 (55.70)	14 (17.72)	21 (26.58)	0.71 (0.86)
104	Unusually loud	27 (34.18)	32 (40.51)	20 (25.32)	0.91 (0.77)
Aggression Problems Syndrome Scale					12.42 (6.74)
3	Argues a lot	35 (44.39)	27 (34.18)	17 (21.52)	0.77 (0.78)
16	Cruelty, bullying, or meanness	57 (72.15)	13 (16.46)	9 (11.39)	0.39 (0.69)
19	Demands a lot of attention	13 (16.46)	26 (32.91)	40 (50.63)	1.34 (0.75)
20	Destroys his/her own things	39 (49.37)	17 (21.52)	23 (29.11)	0.80 (0.87)
21	Destroys others' things	50 (63.29)	13 (16.46)	16 (20.25)	0.57 (0.81)
22	Disobedient at home	32 (40.51)	33 (41.77)	14 (17.72)	0.77 (0.73)
23	Disobedient at school	56 (70.89)	17 (21.52)	6 (7.59)	0.37 (0.62)
37	Gets in many fights	54 (68.35)	14 (17.72)	11 (13.92)	0.46 (0.73)
57	Physically attacks people	47 (59.49)	22 (27.85)	9 (11.39)	0.51 (0.70)
68	Screams a lot	30 (37.97)	30 (37.97)	19 (24.05)	0.86 (0.78)
86	Stubborn, sullen, or irritable	11 (13.92)	37 (46.84)	31 (39.24)	1.25 (0.69)
87	Sudden changes in mood	24 (30.38)	30 (37.97)	25 (31.65)	1.01 (0.79)
88	Sulks a lot	46 (58.23)	21 (26.58)	12 (15.19)	0.57 (0.75)
89	Suspicious	62 (78.48)	14 (17.72)	3 (3.80)	0.25 (0.52)
94	Teases a lot	52 (65.82)	20 (25.32)	7 (8.86)	0.43 (0.65)
95	Temper tantrums or hot temper	26 (32.91)	25 (31.65)	28 (35.44)	1.03 (0.83)
97	Threatens people	71 (89.87)	6 (7.59)	2 (2.53)	0.13 (0.40)
104	Unusually loud	27 (34.18)	32 (40.51)	20 (25.32)	0.91 (0.77)
Withdrawn/Depressed Syndrome Scale					4.22 (3.16)
5	There is very little he/she enjoys	40 (50.63)	16 (20.25)	23 (29.11)	0.78 (0.87)
42	Would rather be alone	27 (34.18)	25 (31.65)	27 (34.18)	1.00 (0.83)
65	Refuses to talk	43 (54.43)	23 (29.11)	13 (16.46)	0.62 (0.76)
69	Secretive, keeps things to self	61 (77.22)	14 (17.72)	4 (5.06)	0.28 (0.55)
75	Too shy or timid	45 (56.96)	24 (30.38)	10 (12.66)	0.56 (0.71)
102	Underactive, slow moving	68 (86.08)	8 (10.13)	2 (2.53)	0.15 (0.43)
103	Unhappy, sad, or depressed	58 (73.42)	19 (24.05)	2 (2.53)	0.29 (0.51)
111	Withdrawn	44 (55.70)	28 (35.44)	7 (8.86)	0.53 (0.66)

Note. CBCL/6-18 = Child Behaviour Checklist for Ages 6-18. Response categories “0”, “1”, “2” are labelled as “Not true”, “Somewhat or sometimes true” and “Very or often true” respectively. Some wordings of the CBCL/6-18 items are abbreviated.

In the fifth and final chapter of this dissertation, I discuss and synthesise the findings of the two studies reported in Chapters 2 and 4. I also extend on the limitations of each study, as well as the specific implications and significance of the findings.

CHAPTER 5

Discussion

Behavioural tools, such as the Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2000, 2001) and the Swanson, Nolan, and Pelham ADHD Rating Scale (SNAP-IV; Swanson et al., 2012), are used in clinical settings to screen for emotional and behavioural problems in children and adolescents, and to monitor their behaviour over time. Information gathered from such tools are of clinical significance, and therefore should be accurate and trustworthy. Tool validity and reliability are inextricably tied to the context in which a tool is administered (Dima, 2018). Factors such as language, sociocultural background, and phenotype of the respondent or subject may affect the validity of tool items (De Kock et al., 2013).

In South Africa, clinicians and researchers frequently use parent-report tools that are generally designed for direct application to English-speaking populations in the global North. Many of these tools have not been formally validated for the linguistically and culturally diverse populations in South Africa. Likewise, many tools measuring child cognition and behaviour are designed with typically developing children in mind. These tools may not be suitable for use with children who have neurodevelopmental disorders (NDDs). Without sufficient evidence to support the adequacy and appropriateness of such tools in these contexts, the dependability of their results remains doubtful. Comprehensive psychometric analyses are needed to arrive at meaningful and accurate conclusions about the measurement properties of the tools in question.

The overarching aim of this dissertation was to investigate the psychometric properties of two behavioural tools; one, the Achenbach System of Empirically Based Assessment (ASEBA) CBCL and Youth Self-Report (YSR), a family of screening tools widely used in many different cultural contexts, and the other, the SNAP-IV, an Attention-Deficit/Hyperactivity Disorder (ADHD) rating scale frequently administered to parents of children with NDDs. This dissertation comprised two distinct studies corresponding to the two behavioural tools.

The two studies presented in the body of this dissertation addressed these aims directly. I summarise the findings of the two studies separately below. Following this, I discuss additional limitations of each study and the significance of the findings overall. I conclude by recommending directions for future research.

Systematic Review of the Psychometric Properties of the ASEBA Forms in Sub-Saharan Africa (SSA)

The primary aim of the systematic review was to investigate the psychometric properties of the ASEBA forms in sub-Saharan African (SSAn) countries, where translated versions of the forms are frequently administered. We identified 145 studies, mostly from South Africa and East African countries, that used the ASEBA forms to measure child behavioural and emotional problems. However, only 58 of those studies reported any measurement properties of the ASEBA forms, and the majority reported only rudimentary psychometric statistics (usually coefficient alpha as a measure of internal consistency). We also evaluated eight psychometric-focused studies against rigorous Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines for good measurement properties (Mokkink et al., 2018; Terwee et al., 2018). All psychometric-focused studies hailed from East African countries. For both the CBCL and the YSR, structural validity was the only measurement property with consistently sufficient results across studies. This promising finding suggests that the established latent structure of the ASEBA subscales, supported by data from the United States, is similar in SSAn populations. The "risk of bias" ratings were typically good for measurement properties with fewer standards, such as internal consistency. Conversely, measurement properties with more stringent standards, such as reliability, were often of inadequate quality. Overall, evidence to support the validity and reliability of the ASEBA forms in SSAn countries is currently limited, both in terms of the quantity and quality of available data.

There were various limitations to interpretation and generalisability of the findings. First, most studies that administered a translated version of an ASEBA form did not include details vis-à-vis the translation procedures. Hence, we were not able to evaluate the quality of the translations across studies. As discussed in previous chapters, a poor quality translation may introduce measurement bias and compromise a tool's validity (De Kock et al., 2013; Van Widenfelt et al., 2005). Thorough evaluation of a translation therefore serves as a prerequisite to further psychometric investigation. In other words, without sufficient knowledge of the qualitative validity of a translation, interpretation of its psychometric properties (e.g., structural validity and internal consistency) is moot. Although this issue was not within the scope of the present review, evaluating a translation process is an additional but necessary layer of psychometric investigation and warrants greater attention in the literature. The findings of this systematic review should be interpreted with this limitation in mind. Second, it goes without saying that the conclusions may not generalise to all

behavioural assessment tools and populations within SSA. For example, it was particularly interesting to note that although most of the included studies hailed from South Africa, none reported psychometric analyses beyond internal consistency. Notwithstanding these limitations, the systematic review serves as a useful framework for clinicians and researchers in SSA to critically evaluate the measurement properties of available assessment tools and select the most appropriate tool based on the available evidence.

Validity of the SNAP-IV in a Sample of Children with NDDs

This study aimed to confirm the structural and construct validity of the SNAP-IV in a sample of children with NDDs in South Africa. A confirmatory factor analysis (CFA) tentatively supported the two-factor structure of the SNAP-IV ('Inattention' and 'Hyperactivity-Impulsivity'), with the exception of three items that did not seem to measure either factor well. Items 15 ("Talks excessively") and 16 ("Blurts out answers") were likely indicators of a child's verbal abilities, rather than a measure of hyperactive/impulsive behaviour. When these problematic items were excluded, the CFA goodness of fit indices improved. In addition, the subscales (sans the problematic items) had good internal consistencies, as measured by ordinal coefficient alpha. The SNAP-IV did not demonstrate convergent and discriminant validity when correlating ADHD-related behaviours with internalising and externalising behaviours respectively, as measured by the CBCL for Ages 6-18 (CBCL/6-18). This result was also likely a consequence of NDD-related factors that affected responses to items on the CBCL/6-18. The SNAP-IV may therefore be less useful as a measure of ADHD for children with NDDs who are non-verbal. The study was underpowered given the smaller sample size ($N = 109$). Notwithstanding this limitation, the data illustrate that the clinical context of tool administration may have important implications for the validity of items and the interpretation of responses.

Although not the focus of the study, it is possible that linguistic and cultural biases may also have influenced the validity of the SNAP-IV scores in this sample. The NeuroDEV South Africa study team made every effort to ensure conceptual equivalence of the items during the review process. However, even "good" translations may have limitations in practice. Grobler (2009) studied the subjective experiences of research assistants administering questionnaires, including the CBCL, that had been translated into various African languages. The research assistants reported that the translations, although technically correct, were often too "heavy", did not "flow" well, and did not align with the colloquial language spoken by participants on a daily basis. As such, many participants found it difficult to understand the translated items, which resulted in research assistants switching between

language versions or translating "freely" in the moment. In addition, the research assistants often needed to rephrase items or provide additional examples to aid comprehension.

We experienced similar challenges when administering the Xhosa-translated versions of the SNAP-IV and CBCL/6-18 to study participants. A Xhosa-speaking research nurse and I, at times, had to rephrase items, simplify the wording of certain items, or switch between the Xhosa and English versions. One example of a particularly challenging SNAP-IV item was Item 14 (How often is your child "on the go" as if "driven by a motor"?). The phrase "on the go" was translated as *esendleleni*, which means "on the road" or "on the way". The different connotations of the English and Xhosa phrasing may have resulted in confusion or misunderstanding of the item. In addition, parents did not seem to understand the concept of a child behaving as if "driven by a motor" (*-qhutywa yinjini*). It seems that this expression did not translate well into Xhosa, which may have compromised the true meaning of the question. An example of a good, conceptually equivalent translation was that of Item 8 ("How often is your child distracted by extraneous stimuli?"). In this instance, however, the difficulty level of the Xhosa translation (*Usoloko ephazamiseka ngezinye izinto ezenzekayo...?*) was much lower than that of the original English wording. It is therefore possible that the language of administration may have influenced the validity of scores on the SNAP-IV. The sample size of the current study was also too small to conduct multigroup factor analyses for the different language groups. A larger, stratified sample would have allowed us to compare the structural validity of the SNAP-IV by language of administration. Our practical experiences in administering the Xhosa-translated tools also highlighted the importance of developing translations that are not only conceptually equivalent but that also consider subtler concerns, such as differential item difficulty and familiarity of expressions and ideas in the target population. In addition, our experiences underscore the value of pilot testing translations in representative samples to detect and counteract such issues prior to the commencement of data collection.

Conclusion and Recommendations for Future Research

Psychometrics is an important but often neglected area of research in SSA. Comprehensive psychometric evaluations inform whether a tool is indeed measuring the construct it purports to measure in a specific population or context. This is of particular importance in South Africa's linguistically and culturally diverse population. Good quality systematic reviews of psychometric studies enable clinicians and researchers to make evidence-based decisions regarding the suitability of a particular tool for a given context. Currently, there are very few systematic reviews of studies that investigate the measurement

properties of behavioural assessment tools in South Africa. This represents a significant gap in the literature and is of clinical significance.

This dissertation addressed two specific aims. They were (i) to conduct a systematic review investigating the use, measurement properties, and cultural appropriateness of the ASEBA forms in SSA, and (ii) to evaluate the validity and internal consistency of the SNAP-IV in the context of children with NDDs. By pursuing these aims, this dissertation made three significant contributions. First, the systematic review showed that while there is some good quality evidence supporting the validity of the ASEBA forms in SSA, the studies investigating their measurement properties are few and inconsistent. This finding will be of value to South African healthcare professionals when considering the ASEBA forms for use in clinical or research settings. Second, this dissertation revealed potential shortcomings of the SNAP-IV apropos its validity in clinical populations. This study will help doctors and psychologists become more aware of possible limitations associated with using standard behavioural tools to measure ADHD-related behaviours in their patients with NDDs. Third, this dissertation highlighted potential problems associated with translating existing behavioural tools into African languages and emphasized the value of conducting thorough translation processes and reporting those procedures in published research. Perhaps most importantly, these findings demonstrated the importance of conducting ongoing psychometric evaluations of widely used behavioural tools and of evaluating the items both statistically and from a qualitative perspective.

More comprehensive psychometric studies with South African samples are needed to arrive at accurate and reliable conclusions about the usefulness of these tools in local contexts. Possible next steps include developing strategies and resources to improve (i) the quality of translation procedures and psychometric analyses and (ii) the reporting of those procedures and analyses in published research. Systematic reviews directly comparing the measurement properties of a variety of commonly used behavioural tools in South Africa would provide a broader overview of the available evidence. In addition, robust studies with larger, more diverse samples of children with NDDs would provide clinicians with further insights into the assessment of comorbid conditions in this population. Taken together, the findings of this dissertation are of practical relevance to South African clinicians and researchers, helping them make informed decisions about the best ways to measure child behaviour in a local context. Finally, this dissertation will hopefully inspire and convince researchers in SSA to prioritise the important work of validating behavioural tools in clinical as well as population-based samples.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. University of Vermont.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. University of Vermont.

De Kock, F., Kanjee, A., & Foxcroft, C. (2013). Cross-cultural test adaptation, translation and tests in multiple languages. In C. Foxcroft & G. Roodt (Eds.), *Introduction to psychological assessment in the South African context* (4th ed., pp. 102–125). Oxford University Press.

Dima, A. L. (2018). Scale validation in applied health research: Tutorial for a 6-step R-based psychometrics protocol. *Health Psychology and Behavioral Medicine*, *6*(1), 136–161. https://doi.org/10.1080/21642850.2018.1472602

Grobler, A. (2009). *Experiences of research assistants in the administration of culturally tailored psychometric data-collection instruments in the Kgolo Mmogo project* [Unpublished masters dissertation]. University of Pretoria. https://repository.up.ac.za/bitstream/handle/2263/26719/dissertation.pdf

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., 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. https://doi.org/10.1007/s11136-017-1765-4

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, C. A., Sergeant, J. A., Clevenger, W., Wasdell, M., McCleary, R., Lakes, K., & Wigal, T. (2012). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of the SNAP and the SWAN Rating Scales. *The International Journal of Educational and Psychological Assessment*, *10*(1), 51–70.

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. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research*, *27*(5), 1159–1170. https://doi.org/10.1007/s11136-018-1829-0

Van Widenfelt, B. M., Treffers, P. D. A., De Beurs, E., Siebelink, B., Koudijs, E., & Siebelink, B. M. (2005). Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child & Family Psychology Review*, *8*(2), 135–147. https://doi.org/10.1007/s10567-005-4752-1

APPENDIX A

Demographic Questionnaire

1. How many languages are spoken in [name]'s home?
   - ☐ 1
   - ☐ 2
   - ☐ 3 or more

2. What is the primary language spoken in [name]'s home?
   - ☐ Afrikaans
   - ☐ English
   - ☐ Xhosa or isiXhosa
   - ☐ Ndebele
   - ☐ Sotho-Northern or Sepedi
   - ☐ Sotho-Southern or Sesotho
   - ☐ Swazi or Swati or siSwati
   - ☐ Tsonga or Xitsonga
   - ☐ Tswana or Setswana
   - ☐ Venda or Tshivenda
   - ☐ Zulu or isiZulu
   - ☐ Unknown
   - ☐ Other. Describe:

3. What is the second language spoken in [name]'s home?
   - ☐ Afrikaans
   - ☐ English
   - ☐ Xhosa or isiXhosa
   - ☐ Ndebele
   - ☐ Sotho-Northern or Sepedi
   - ☐ Sotho-Southern or Sesotho
   - ☐ Swazi or Swati or siSwati
   - ☐ Tsonga or Xitsonga
   - ☐ Tswana or Setswana
   - ☐ Venda or Tshivenda
   - ☐ Zulu or isiZulu
   - ☐ Unknown
   - ☐ Other. Describe:

4. What is the third language spoken in [name]'s home?
   - ☐ Afrikaans
   - ☐ English
   - ☐ Xhosa or isiXhosa
   - ☐ Ndebele
   - ☐ Sotho-Northern or Sepedi
   - ☐ Sotho-Southern or Sesotho
   - ☐ Swazi or Swati or siSwati
   - ☐ Tsonga or Xitsonga
   - ☐ Tswana or Setswana
   - ☐ Venda or Tshivenda
   - ☐ Zulu or isiZulu
   - ☐ Unknown
   - ☐ Other. Describe:

APPENDIX B
Asset Index

1. What is the highest level of education [child's name] has received?
   - ☐ Never attended
   - ☐ 1st grade
   - ☐ 2nd grade
   - ☐ 3rd grade
   - ☐ 4th grade
   - ☐ 5th grade
   - ☐ 6th grade
   - ☐ 7th grade
   - ☐ 8th grade
   - ☐ 9th grade
   - ☐ 10th grade
   - ☐ 11th grade
   - ☐ 12th grade
   - ☐ Non-academic curriculum
   - ☐ Alternative curriculum
   - ☐ Refused to answer
   - ☐ Don't know
   - ☐ Other. Describe:

2. How far did you get in school?
   - ☐ No education
   - ☐ Completed Grade 1 (Sub A) to Grade 5 (Standard 3)
   - ☐ Completed Grade 6 (Standard 4) to Grade 7 (Standard 5)
   - ☐ Completed Grade 8 (Standard 6) to Grade 11 (Standard 9) i.e., high school without matriculating
   - ☐ Completed Grade 12 (Standard 10) i.e., high school with matriculating
   - ☐ Part of university/college/post-matric education
   - ☐ Completed university/ college/post-matric education

3. What is your current employment status?
   - ☐ Working now
   - ☐ Self-employed
   - ☐ Looking for work
   - ☐ Unemployed
   - ☐ Temporarily laid off
   - ☐ Homemaker
   - ☐ Student
   - ☐ Illness/sickness
   - ☐ Disabled
   - ☐ Other. Describe:

4. Do you receive any social assistance in the form of a government grant?
   - ☐ Yes
   - ☐ No

   If yes, what kind of grant do you receive?
   - ☐ Childcare grant
   - ☐ Disability grant
   - ☐ Care dependency grant

5. What is your own average income per month (i.e., average over the past 6 months)?
    - ☐ Less than R1000 per month
    - ☐ R1000-R5000 per month
    - ☐ R5000-R10 000 per month
    - ☐ More than R10 000 per month
    - ☐ Unknown

6. What is your average household income per month (i.e. average over the past 6 months)? Your best estimate is fine.
    - ☐ Less than R1000 per month
    - ☐ R1000-R5000 per month
    - ☐ R5000-R10 000 per month
    - ☐ R10 000-R15 000 per month
    - ☐ More than R15 000 per month
    - ☐ Unknown

7. What is your marital status?
    - ☐ Single (never married)
    - ☐ Not married, but in a marriage like relationship (i.e., living together)
    - ☐ Married
    - ☐ Divorced
    - ☐ Widowed

    How many months have you been in a marriage/marriage like relationship?

    How many times in your life have you been married and/or lived together with a partner (include current partner if living together)?

8. How far did your spouse/partner get in school?
    - ☐ No education
    - ☐ Completed Grade 1 (Sub A) to Grade 5 (Standard 3)
    - ☐ Completed Grade 6 (Standard 4) to Grade 7 (Standard 5)
    - ☐ Completed Grade 8 (Standard 6) to Grade 11 (Standard 9) i.e., high school without matriculating
    - ☐ Completed Grade 12 (Standard 10) i.e., high school with matriculating
    - ☐ Part of university/college/post-matric education
    - ☐ Completed university/college/post-matric education
    - ☐ Unknown

9. What is your spouse/partner's current employment status?
    - ☐ Working now
    - ☐ Self-employed
    - ☐ Looking for work
    - ☐ Unemployed
    - ☐ Temporarily laid off
    - ☐ Homemaker
    - ☐ Student
    - ☐ Illness/sickness
    - ☐ Disabled
    - ☐ Other. Describe:

10. Does your spouse/partner receive assistance in the form of a government grant?
    - ☐ Yes
    - ☐ No

If yes, what kind of grant do you receive?

- ☐ Childcare grant
- ☐ Disability grant
- ☐ Care dependency grant

11. What is your spouse/partner's average income per month (i.e. average of the past 6 months)?
    - ☐ Less than R1000 per month
    - ☐ R1000-R5000 per month
    - ☐ R5000-R10 000 per month
    - ☐ More than R10 000 per month
    - ☐ Don't know

12. Which of the following best describes the area in which you live?
    - ☐ Urban
    - ☐ Rural
    - ☐ Township

13. How many years have you lived at the current address?

14. Which of the following best describes your home (type of dwelling)?
    - ☐ Shack
    - ☐ Wendy house or backyard dwelling
    - ☐ House
    - ☐ Flat
    - ☐ Refugee centre/homeless shelter
    - ☐ Other. Describe:

15. Which of the following do you have in your home?
    - ☐ Electricity
    - ☐ Tap or running water
    - ☐ Domestic servant
    - ☐ A flush toilet inside
    - ☐ A built-in kitchen sink
    - ☐ An electric stove or hotplate
    - ☐ A working telephone (this includes a cell phone)
    - ☐ At least one motor car or truck
    - ☐ A motorcycle or scooter
    - ☐ A bicycle

16. Is the house where you stay owned, rented or an informal settlement plot?
    - ☐ Own
    - ☐ Rent
    - ☐ Neither, informal settlement plot
    - ☐ Other. Describe:

17. Do you own any land other than the land where your house is?
    - ☐ Yes
    - ☐ No

18. Do you personally do any of the following? Check all that apply.
    - ☐ Shop at supermarkets
    - ☐ Use any financial services (such as bank account, ATM card or credit card)
    - ☐ Have an account at a retail store (e.g., Pep, Jet etc)

APPENDIX C

Neuromedical Assessment

1. What is [child's name] sex?
   - ☐ Male
   - ☐ Female

2. What is the current level of [child's name] spoken language?
   - ☐ No spoken language
   - ☐ Single words
   - ☐ Phrases
   - ☐ Fluent

3. What is the highest level of education [child's name] has ever received?
   - ☐ Never attended
   - ☐ Grade 1
   - ☐ Grade 2
   - ☐ Grade 3
   - ☐ Grade 4
   - ☐ Grade 5
   - ☐ Grade 6
   - ☐ Grade 7
   - ☐ Grade 8
   - ☐ Grade 9
   - ☐ Grade 10
   - ☐ Grade 11
   - ☐ Grade 12
   - ☐ Non-academic curriculum
   - ☐ Alternative curriculum
   - ☐ Other

4. DSM-5 neurodevelopmental diagnosis* (select all that apply):
   - ☐ Intellectual Disability
   - ☐ Global Developmental Delay
   - ☐ Communication Disorders
   - ☐ Autism Spectrum Disorders
   - ☐ Attention-Deficit/Hyperactivity Disorder
   - ☐ Specific Learning Disorders
   - ☐ Unspecified neurodevelopmental disorder

*This information was extracted from the child participant's medical records.

APPENDIX D

Swanson, Nolan, and Pelham ADHD Rating Scale (4th ed.; SNAP-IV)

I would like to ask you some questions about your child's behaviour over the past year. For each item, please select the option that best describes your child over the past year.

0 Not at all 1 Just a little 2 Quite a bit 3 Very much NA Not applicable

Inattention

1. How often does [child's name] fail to give close attention to details or make careless mistakes in schoolwork, work, or other activities?
2. How often does [child's name] have difficulty sustaining attention in tasks or play activities?
3. How often does [child's name] not seem to listen when spoken to directly?
4. How often does [child's name] not follow through on instructions and fail to finish schoolwork, chores or duties?
5. How often does [child's name] have difficulty organizing tasks and activities?
6. How often does [child's name] avoid, dislike, or act reluctant to engage in tasks that require sustained mental effort (e.g., schoolwork or homework)?
7. How often does [child's name] lose things necessary for tasks or activities (e.g., toys, assignments, pencils, books, or tools)?
8. How often is [child's name] distracted by extraneous stimuli?
9. How often is [child's name] forgetful in daily activities?

Hyperactivity-Impulsivity

10. How often does [child's name] fidget with their hands or feet or squirm in their seat?
11. How often does [child's name] leave their seat in the classroom or in other situations in which remaining seated is expected?
12. How often does [child's name] run about or climb excessively in situations in which it is inappropriate?
13. How often does [child's name] have difficulty playing or engaging in leisure activities quietly?
14. How often is [child's name] "on the go" or often act as if "driven by a motor"?
15. How often does [child's name] talk excessively?
16. How often does [child's name] blurt out answers before questions have been completed?
17. How often does [child's name] blurt out answers before questions have been completed?
18. How often does [child's name] interrupt or intrude on others (e.g., butts into conversations/games)?

APPENDIX E

Child Behaviour Checklist for Ages 6-18 – Attention Problems Syndrome Scale

Below is a list of items that describe children and youths. For each item that describes your child now or within the past 6 months, please indicate if the item is very/often true, somewhat/sometimes true, or not true (as far as you know) of your child. Please answer all items as well as you can, even if some do not seem to apply to your child.

0 Not true (as far as you know) 1 Somewhat or sometimes true 2 Very or often true

Item Number	Item
1	Acts too young for his/her age
4	Fails to finish things he/she starts
8	Can't concentrate, can't pay attention for long
10	Can't sit still, restless, or hyperactive
13	Confused or seems to be in a fog
17	Daydreams or gets lost in his/her thoughts
41	Impulsive or acts without thinking
61	Poor school work
78	Inattentive or easily distracted
80	Stares blankly

APPENDIX F

Child Behaviour Checklist for Ages 6-18 – ADHD DSM-Oriented Scale

Below is a list of items that describe children and youths. For each item that describes your child now or within the past 6 months, please indicate if the item is very/often true, somewhat/sometimes true, or not true (as far as you know) of your child. Please answer all items as well as you can, even if some do not seem to apply to your child.

0 Not true (as far as you know) 1 Somewhat or sometimes true 2 Very or often true

Item Number	Item
4	Fails to finish things he/she starts
8	Can't concentrate, can't pay attention for long
10	Can't sit still, restless, or hyperactive
41	Impulsive or acts without thinking
78	Inattentive or easily distracted
93	Talks too much
104	Unusually loud

APPENDIX G

Child Behaviour Checklist for Ages 6-18 – Aggressive Behaviour Syndrome Scale

Below is a list of items that describe children and youths. For each item that describes your child now or within the past 6 months, please indicate if the item is very/often true, somewhat/sometimes true, or not true (as far as you know) of your child. Please answer all items as well as you can, even if some do not seem to apply to your child.

0 Not true (as far as you know) 1 Somewhat or sometimes true 2 Very or often true

Item Number	Item
3	Argues a lot
16	Cruelty, bullying, or meanness to others
19	Demands a lot of attention
20	Destroys his/her own things
21	Destroys things belonging to his/her family or others
22	Disobedient at home
23	Disobedient at school
37	Gets in many fights
57	Physically attacks people
68	Screams a lot
86	Stubborn, sullen, or irritable
87	Sudden changes in mood or feelings
88	Sulks a lot
89	Suspicious
94	Teases a lot
95	Temper tantrums or hot temper
97	Threatens people
104	Unusually loud

APPENDIX H

Child Behaviour Checklist for Ages 6-18 – Withdrawn/Depressed Syndrome Scale

Below is a list of items that describe children and youths. For each item that describes your child now or within the past 6 months, please indicate if the item is very/often true, somewhat/sometimes true, or not true (as far as you know) of your child. Please answer all items as well as you can, even if some do not seem to apply to your child.

0 Not true (as far as you know) 1 Somewhat or sometimes true 2 Very or often true

Item Number	Item
5	There is very little he/she enjoys
42	Would rather be alone than with others
65	Refuses to talk
69	Secretive, keeps things to self
75	Too shy or timid
102	Underactive, slow moving, or lacks energy
103	Unhappy, sad, or depressed
111	Withdrawn, doesn't get involved with others

APPENDIX I

UCT Human Research Ethics Committee – Protocol Approval Letter

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E53-46 Old Main Buildin
Groote Schuur Hospita
Observatory 792
Telephone [021] 406 649
Email: sumayah.ariefdien@uct.ac.z
Website: www.health.uct.ac.za/fhs/research/humanethics/form

26 June 2019

HREC REF: 357/2019

A/Prof K Donald
Paediatric Neurodevelopment
4th floor
Institute of Child Health Building
Red Cross War Memorial Children's Hospital
Rondebosch

Dear A/Prof Donald

PROJECT TITLE: EVALUATING THE PSYCHOMETRIC PROPERTIES OF NEURODEVELOPMENTAL ASSESSMENT TOOLS IN A SOUTH AFRICAN CONTEXT (SUB-STUDY LINKED TO 810/2016) MMED CANDIDATE - MS M ZIEFF

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee (HREC) for review.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30 June 2020.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

We acknowledge that the student: Ms M Zieff will also be involved in this study.

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval, where necessary, before the research may occur.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE

APPENDIX J

UCT Human Research Ethics Committee – Protocol Amendment Approval Letter

Form FHS006: Protocol Amendment

HREC office use only (FWA00001637; IRB00001938)		
Approved	Type of review: Expedited	Full committee
This serves as notification that all changes and documentation described below are approved.		
Signature Chairperson of the HREC	Date	26/2/20

Note: All major amendments must include a local **PI Synopsis** justifying the changes for the amendment. Please note that incomplete amendment submissions will not be reviewed.

Comments from the HREC to the Principal Investigator:
Note: The approval of this protocol amendment does not grant annual approval. Please complete the FHS016 / FHS017 form for annual approval at least one month before study expiration.

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	
HREC REF Number	367/2019
Protocol title	Evaluating the psychometric properties of neurodevelopmental assessment tools in a South African context (sub-study linked to 810/2016)
Protocol number (if applicable)	
Principal Investigator	Prof Kirsten Donald
Department / Office Internal Mail Address	Department of Paediatrics and Child Health

1.1 Is this a major or a minor amendment? (see FHS006hlp) Major (tick box) Minor (tick box)	Major	Minor
1.2 Does this protocol receive US Federal funding?	Yes	No
1.3 If the amendment is a major amendment and receives US Federal Funding, does the amendment require full committee approval? **Note:** Any protocol amendments for **Full Committee** review MUST be submitted on the monthly HREC submission dates. (Please email an electronic copy to hrec-enquiries@uct.ac.za)	Yes	No

Note: Only the first page is included here.

APPENDIX K

UCT Human Research Ethics Committee – Annual Progress Report

UNIVERSITY OF CAPE TOWN

FACULTY OF HEALTH SCIENCES
Human Research Ethics Committee

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30.6.22
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee		Date Signed	23/7/21

Note: Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za.
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website:
http://www.health.uct.ac.za/fhs/research/humanethics/forms

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	22 July 2021		
HREC REF Number	367/2019	Current Ethics Approval was granted until	30 June 2021
Protocol title	Evaluating the psychometric properties of neurodevelopmental assessment tools in a South African context		
Protocol number (if applicable)			
Are there any sub-studies linked to this study?		☐ Yes	✓ No
If yes, could you please provide the HREC Reference number for all sub-studies? **Note:** A separate FHS016 must be submitted for each sub-study.	This is a sub-study linked to 810/2016.		
Principal Investigator	Professor K Donald		
Department / Office Internal Mail Address	Department of Paediatrics and Child Health Red Cross War Memorial Children's Hospital Rondebosch		

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Note: Only the first page of the most recent progress report is included here.

APPENDIX L

UCT Human Research Ethics Committee – NeuroDEV South Africa Protocol Approval Letter

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E52-24 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 404 7682 • **Facsimile** [021] 406 6411
Email: nosi.tsama@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

08 March 2017

HREC REF: 810/2016

A/Prof K Donald
Paediatric Medicine
Paediatrics- SCAH
Rondebosch

Dear A/Prof Donald

PROJECT TITLE: NEURO-DEVELOPMENTAL EXOME VARIANTS IN AFRICAN POPULATIONS (NEURODEV)

Thank you for submitting your response letter to the Faculty of Health Sciences Human Research Ethics Committee dated 20 February 2017.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30th March 2018.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval before the research may occur.

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical

APPENDIX M

Western Cape Government – Approval to Collect Data

TYGERBERG HOSPITAL
REFERENCE:
Research Projects
ENQUIRIES: **Dr GG**
Marinus
TELEPHONE:**021 938 5752**

Ethics Reference: **HREC REF: 810/2016**

TITLE: **NEURO-DEVELOPMENTAL EXOME VARIANTS IN AFRICAN POPULATIONS [NEURODEV]**

Dear Prof Donald

PERMISSION TO CONDUCT YOUR RESEARCH AT TYGERBERG HOSPITAL.

1. In accordance with the Provincial Research Policy and Tygerberg Hospital Notice No 40/2009, permission is hereby granted for you to conduct the above-mentioned research here at Tygerberg Hospital.

2. Researchers, in accessing Provincial health facilities, are expressing consent to provide the Department with an electronic copy of the final feedback within six months of completion of research. This can be submitted to the Provincial Research Co-Ordinator (Health.Research@westerncape.gov.za).

DR GG MARINUS
MANAGER: MEDICAL SERVICES

DR D ERASMUS
CHIEF EXECUTIVE OFFICER
Date: 21 May 2018

Administration Building, Francie van Zijl Avenue, Parow, 7500 Private Bag X3, Tygerberg, 7505
tel: +27 21 938-6267 fax: +27 21 938-4890 www.capegateway.go.v.za

Western Cape Government
Health

DR AN PARBHOO
Manager: Medical Services
Red Cross War Memorial Children's Hospital
Email: Anita.Parbhoo@westerncape.gov.za
Tel: +27 21 658 5430 Fax: +27 21 658 5006/5166

08 August 2018

Prof K Donald
Paediatric Medicine

Dear Prof Donald,

RESEARCH: RXH: RCC 143

PROJECT TITLE: Neuro-developmental Exome Variants in African Populations (Neurodev)

It is a pleasure to inform you that approval is hereby granted to conduct above-mentioned study at Red Cross War Memorial Children's Hospital,

Yours sincerely,

DR AN PARBHOO
MANAGER: MEDICAL SERVICES

APPENDIX N

Harvard T.H. Chan School of Public Health IRB – Ethical Approval

HARVARD
Human Research Protection Program

Harvard T.H. Chan School of Public Health
Office of Human Research Administration
90 Smith Street, 3rd Floor
Boston, MA 02120
Federalwide Assurance FWA00002642

Notification of Initial Study Approval

September 28, 2017

Elise Robinson
erobinso@hsph.harvard.edu

Protocol Title:	Genetic Characterization of Neuro-Developmental Disorders in South African Populations
Principal Investigator:	Elise Robinson
Protocol #:	IRB17-1260
Funding Source:	Stanley Center [Broad]
Review Date:	9/13/2017
STUDY Effective Date:	9/13/2017
Expiration Date:	9/12/2018
IRB Review Type:	Full Committee
IRB Review Action:	Approved

The Institutional Review Board (IRB) of the Harvard T.H. Chan School of Public Health approved this Initial Study. **Please note that the approval for this protocol will lapse on 9/12/2018.**

This approval includes the following:

- Initial Application, IRB17-1260
- IRB Protocol: IRB17-1260 Elise Robinson NeuroDEV UCT protocol 09.01.17 - clean.docx (1)
- Recruitment Materials: NeuroDev SA - Recruitment script - clean.pdf (1)
- Recruitment Materials: Control screener - Neurodevelopmental Screening Tool.pdf (1)
- Consent Form: NeuroDEV UCT Assent - case & control - 09.01.17 clean.pdf (1)
- Consent Form: NeuroDEV UCT Info & Consent - parent of case - 09.28.17 clean.pdf (1)
- Consent Form: NeuroDEV UCT Info & Consent - parent of control - 09.28.17 clean.pdf (1)
- Consent Form: NIH Atlas_Photo Consent for Cases - 8.10.17.pdf (1)
- Consent Form: UBACC adapted for NeuroDev 08.10.17.pdf (1)
- Study Instrument/Tools: NeuroDEV South Africa Phenotype Tools - Parents - 07.06.17.pdf (1)
- Study Instrument/Tools: NeuroDEV South Africa Phenotype Tools - Child controls - 08.10.17.pdf (1)

Note: Only the first page is included here.

APPENDIX O

NeuroDev South Africa information and consent form parent/guardian of case child

You and your child are invited to take part in a research study. Please take some time to read the information here. Please ask questions about any part of this project that you do not understand. It is very important that you feel you understand what this research is about and how you could be involved. Also, being in the study is voluntary**.** You are free to choose not to join. If you say no, this will not affect you in any way. This study has been approved by the University of Cape Town Human Research Ethics Committee (HREC ref 801/2016) and the Harvard School of Public Health, in the United States of America. The study will be carried out according to the highest ethical standards.

What is the study about?

This study looks at development in childhood. The study will try to find out about genes that may be related to why some children develop differently than others. We will be asking over 5,000 people to take part in this study. The reason is to better understand the cause of delays in development and other social and thought problems.

Background

Our body is made up of many tiny parts called cells. These cells contain DNA, which tell the cells how to grow and work. We get our DNA from our parents. Nobody else in the world has the same DNA as you, unless you are a twin. Difference in DNA can be why some people are tall and others are short. Some illnesses are caused by problems with DNA. The study of DNA is called genetics. DNA can be found in blood and in saliva. A cell line is a culture made from cells of the same type. Under the right conditions, these cells can make copies of themselves. So without having to draw blood again, researchers can keep studying the cells in your blood for many years. This makes cell lines very useful. A stem cell is a special type of cell. Stem cells can form many different types of cells in the body. Stem cells are made of the same DNA as all other cells in your body. A stem cell made from your blood is called an induced pluripotent stem cell (iPS cell).

Do I have to take part?

It is your choice to be in this study. This study does not have anything to do with the care you or your child receives at this hospital or school. If you do not want to answer some questions or if you do not want some tests to be done, you can tell us. If you would like to stop at any time, you can. You can say "Yes" and join the study; or you can also say "No", you don't want to join.

What does being in this study include?

What your child will do

If you choose to join your child in this study, we will ask you some questions about your child to assess your child's behavior, reasoning, and movements. We will also examine your child's body and look at your child's medical records to find out about his or her health. The whole process could take up to 3 hours. We would be happy to provide the names of the tests and explain these tests to you before we start. A nurse will ask your child to give some blood. Your child will give up to two teaspoons (10 ml) of blood. If you agree, we may also ask for your child to give us some saliva in a small plastic tube. If we collect saliva, we will ask for less than half a teaspoon (2 ml or less) of saliva. DNA is in blood and saliva, which we call a 'sample'.

The sample will be sent for DNA processing in the United States of America. We will analyse your child's DNA using chemicals and computers. Then we will compare your child's DNA to the DNA of other children in the study to look for differences. No one testing your child's DNA will know your child's name. You also have the choice of contributing your child's sample for cell line and stem cell research. Stem cells are important because they can become all types of cells in the body, such as muscle cells, brain cells or heart cells. Stem cells also can make copies of themselves, and for this reason a stem cell line growing in a lab could exist for a very long time, even longer than your child's lifetime. With your permission, we will also photograph your child's face, hands and feet. The photographs are to help doctors and other health experts recognize children with or without. Your child does not have to be photographed to be in the study. If you agree to have your child photographed, the photos will be used for clinical, educational and research purposes. You can also choose to allow public use of the photo to help improve diagnosis

of children with difficulties like your child.

What you will do

You are also invited to enrol yourself in addition to your child. If you accept, you will give about three teaspoons (14,5 ml) of blood, which contains your DNA. In addition, we will ask you to do a puzzle that looks at how you reason. In the same way as for your child, your blood will be sent for DNA processing in the United States of America. Then we will compare your DNA to your child's DNA to look for differences. No one analysing your or your child's DNA will know your names. If you agree to be part of the cell line research, we may make cell lines from your blood, and we may try to make stem cells.

Storage and sharing of your information and samples

With your permission, your and your child's blood sample and health information will be sent to and stored in the United States of America. Some of the sample may also be stored here in South Africa. Your child's donated DNA can last a very long time. The samples will be stored and may be used for other tests in the future. Access to your child's DNA will continue to be guarded into the future by a set of strict controls to make sure that it is only used to study questions that are important to science.

Creating a cell 'bank' or library

The samples will be stored at the National Institutes of Mental Health Repository and Genomics Resource (NIMH-RGR). This is a 'bank' or library for biological samples funded by the department of health in the United States government. Some of the samples will also be stored in South Africa at the Department of Human Genetics in the University of Cape Town. That way, your and your child's samples can be shared with other researchers at universities, hospitals, government agencies and companies around the world. One place your DNA will be shared with is the Broad Institute, a non-profit American research institute with skilled scientists who know how to analyse DNA. If you agree to be part of the cell line research, the cells and your health information will be stored in the cell library and distributed for a long time, possibly forever. None of the researchers who study your cells or health information will know your name. The samples and information will be available to scientists all around the world, to study any research question. Examples include research to understand the cause of diseases (for example heart disease, cancer, or psychiatric disorders). Other examples include research to develop new methods, to study human behaviours, or to find out where different groups of people come from.

If you agree to participate in the cell line and stem cell research, your samples could be of even greater use. Researchers around the world may use your cells to study the biology of stem cells, certain illnesses, and whether it is possible to transplant stem cells (or things made by those cells) to treat diseases. Common examples of what might happen to the stem cells include: Testing the cells' DNA, using cells to test drugs to treat disease, changing some of the genetic code within the cells, transplanting the cells (or things made by those cells) to humans or animals, sharing cells widely for research, training or to make commercial medical products. Because these samples can last forever, scientists may come up with new ways of working with them in the future. Right now, we do not know what the future uses may be.

Creating a photo library

If you allow us to photograph your child, we will ask you to choose one or both of the following options:

1. The photos can be shared on a mobile application that will improve technology for diagnosing children with delays in development in the future. One example of a mobile application that uses photos to help diagnose genetic conditions is Face2Gene. If a photo is shared this way, a computer breaks it up into thousands of bits of information and compares it other photos. When the photo is broken up in this way, it cannot be used to identify your child.
2. The photos can also be shared on a public website called the NIH Atlas of Human Malformations in Diverse Populations. The NIH Atlas is an online photo library that can be used by doctors to help diagnose children now. The website is made for doctors and health professionals, but anyone can see it.

If you agree to one or both of these uses of your child's photos, they will be shared together with some pieces of health information (e.g., age, ethnicity & diagnosis).

Putting information in public databases

Some of the genetic and health information from this study will be freely available in public databases that anyone can use. Examples of these databases are ClinVar and the Genome Aggregation Database (GnomAD). These public open-access databases may include information about your DNA, as well as your ethnic group and sex. This public information will not be labeled with your name or other information that could easily identify you.

Your and your child's DNA and health information will also be put in one or more controlled-access database. This means that only researchers who apply for and get permission to use the information for a specific research project will be able to access it. Your and your child's DNA and health information will not be labelled with your names or other information that could be used to identify you.

Securing information on a group of computers

Your and your child's DNA and health information will not be labelled with your names or other information that could be used to identify you. To ensure safety, the information will be stored and analysed in the Cloud. "The Cloud" refers to a group of computers in a data centre that is accessible through the Internet. The Cloud is managed by a company or other institution that is not part of the research team. The study information will be stored on a secure Cloud server to prevent unwanted access. The company or other institution that will process and store your data is located in the United States. It may also come to be located in other countries where the laws may not protect your privacy to the same level as in your country. However, all reasonable steps will be taken to protect your privacy.

What if I change my mind?

If you join the study, you can change your mind later and decide that you don't want to be part of it anymore, and you do not want your child's blood or saliva to be used. Please let us know and we will destroy the sample. If your or your child's sample has already been tested at the time you change your mind, your results and other data may have already been shared with other investigators. In that case, we will not be able to destroy this data. However, the code that links your child's name to their information and samples can be destroyed. Whether or not you decide to join the study, the way we look after you and your child in the clinic or at school will be the same. It is your choice whether to be in the study or not.

Benefits and Risks

Benefits

For most people, there are no direct benefits to being in this study. However, in some cases we may find something in your child's DNA that helps explain what caused his or her condition. If we find something, we would have to double check it. That would mean inviting you back to give more blood and say if you still want to get the results. In most cases, we will not find results to return. Not finding a result does not mean there is no genetic cause, but only that we do not yet know what the genetic cause(s) might be. One reason somebody might want to share their DNA without direct benefit is that it may help scientists to learn more about the human body, and better understand problems of child development. We hope this will help make better treatments in the future.

Risks

The risks of taking blood include pain, as with a pin prick, a bruise at the point where the blood is taken, and redness and swelling of the vein. There is also a rare risk of infection, if the area is not kept clean, and a rare risk of fainting. When answering questions about yourself or your child, you may experience distress or discomfort. You do not have to answer questions if they make you uncomfortable. You may also ask to take a break at any time. If you agree to have your child photographed and to share the photo publicly, it is possible that somebody somewhere - perhaps, somebody who looked after your child in hospital, or a friend or relative - may identify your child. There are some risks related to the use of your or your child's information and samples. It is possible that someone could break into the system that stores information about you and your child. All the computers with study results are protected by several strong passwords, so this is very unlikely. It is also possible that your child could be identified through the photographs, which will be stored on the Cloud. Security of the Cloud is high, so breaking into the Cloud is very unlikely. There may be other privacy risks that we do not know. If someone has another DNA sample from

you or your child, which has not been protected, it is possible that you could be identified from your sample. This would happen if the two samples were matched, and one of them was linked to your name. The risk of this happening is very small but may grow in the future as genetic processing is more widely used. As certain conditions and traits run in families and are inherited. You may learn something about your health that relates to the health of your relatives. Although your DNA is unique to you, you share some similarities with your children, parents, brothers, sisters, and other blood relatives. Therefore, finding out about your results could mean you learn something that relates to the health of your family members and might cause you or your family distress. Since some genetic information can help predict future health problems for you and your relatives, this information might be of interest to health care providers, insurance companies and others. However, South African law and policy provides some protections against discrimination on the basis of an individual's "state of health" (e.g., the South African Medical Schemes Act No. 131 of 1998).

Privacy

The information we get from you and your child will not be shared with your health insurance. Nobody outside the study will be able to know which results belong to you or your child. The researchers will not tell anyone else that either of you are in the study.

Reimbursement

You will be reimbursed for your time, travel costs and the inconvenience of taking part in this study. Once you have been interviewed and you and your child have given your blood, and possibly saliva, the study staff will reimburse you. You will be reimbursed 250 Rand. As a thank you, we will also give your child a goodie bag, which may include a small toy, snacks and a polaroid photo of your family.

The control over the samples you and your child donate will be held by the University of Cape Town and The Broad Institute. Your and your child's samples will not be sold, but investigators may make products based on studying these and other samples, images or information. If this happens, you will not be able to share in any profits. This is not likely to happen for many years, if at all.

Future Contact

Individual results from this research project will not be given back to you or put into your or your child's medical records, but we may learn something from this study that we would like to follow up on in the future. We may also want to contact you in the future to see if you or your child would be interested in enrolling in another study. Please indicate on the consent form if you are willing to be contacted in the future. You have the right to withdraw your consent for future contact at any point in the future.

When your child turns 18, he/she has the right to decide whether to be in research and the right to withdraw from this study. We will not get back in touch when your child turns 18, unless you or your child seeks to do so. However, if your child reaches the age of 18 during the course of the study, we will contact you and your child, and offer your child consent. If your child does not have the capacity to consent, we will offer you to re-consent. If we cannot reach you, or your child, or you or your child chooses not to consent, we will break all links to their identity.

Questions

If you have any questions about this study, please ask the research nurse. You can also contact **the Project Manager** emma.eastman@uct.ac.za **and the Primary Investigator** Kirsty.donald@uct.ac.za; **081 810 2614** to have your questions answered. If you wish to withdraw your consent later, you may also contact the project manager. For questions about your rights as a research participant, please contact a representative of the **University of Cape Town Human Research Ethics Committee on 021 406 6338**. This can include questions about your participation, complaints or concerns about the study, including if you feel pressure to enrol.

You will be given a copy of this information sheet and signature page for your records.

Please keep a copy of these documents in case you want to read them again.

Consent Form: Parent/Guardian of Control

I, _____the legal guardian of _____agree to allow my child to take part in this study, which will use his/her DNA to try to learn about developmental delays and other social and thought problems.

- ☐ I give permission to be interviewed and for my answers to be recorded.
- ☐ I give permission for a sample of my child's blood and saliva to be taken.
- ☐ I give permission for my child's sample to be stored at the Human Genetics Lab at the University of Cape Town, at the National Institute of Mental Health Repository and Genomics Resource in the United States and at the Broad Institute in the United States.
- ☐ I agree to my child's DNA being made available on both open and controlled access databases.

I understand that:

- ☐ My child's identity will be protected.
- ☐ I will be compensated 250 Rand for the time and inconvenience of being part of the study.

Additionally, I am choosing to give permission for the following:

1. Do we have your permission to create cell lines and stem cells from the blood we draw?

☐ Yes
☐ No

2. Do we have your permission to photograph your child?

☐ Yes, but the photo can only be shared outside of the study with health care providers or researchers for clinical, educational or research purposes (for example with the Face2Gene application)
☐ No

3. Do we have your permission to contact you in the future to follow-up on this study or for future studies?

☐ Yes
☐ No

Address:

_____

Telephone/Cell:

_____

Email/Other:

_____

4. If you agree to be contacted again, and we cannot find you, may we try to reach you through a friend or relative?

☐ Yes
☐ No

Contact information of a relative/friend (1):

Name: _____ Relationship: _____

Telephone/Cell: _____ Email/Other: _____

Contact information of a relative/friend (2):

Name: _____ Relationship: _____

Telephone/Cell: _____ Email/Other: _____

I have read, or have been read, the accompanying information sheet in my own language. I understand this consent form and the information sheet. Any questions I had have been answered. I understand that I may withdraw my consent from any, or all, aspects of this study without any impact on my medical care, and without any repercussions.

Signed in _____(place) on _____(day) of _____(month) _____ (year)

Signature _____

Witness 1 Name _____Signature _____