

Other

Original Research Paper

Cite this article: Ceccarelli C *et al* (2021). Validation of the Communication Profile-Adapted in Ethiopian children with neurodevelopmental disorders. *Global Mental Health* 8, e45, 1–11. <https://doi.org/10.1017/gmh.2021.44>

Received: 9 April 2021

Revised: 21 October 2021

Accepted: 11 November 2021

Key words:


Caregiver; communication disorders; neurodevelopmental disorders; psychometrics; sub-Saharan Africa

Author for correspondence:

Rosa A. Hoekstra,

E-mail: rosa.hoekstra@kcl.ac.uk

Validation of the Communication Profile-Adapted in Ethiopian children with neurodevelopmental disorders

Caterina Ceccarelli¹, Ioannis Bakolis^{2,3}, Bethlehem Tekola¹, Mersha Kinfe⁴, Anton Borissov¹, Fikirte Girma⁴, Rehana Abdurahman⁵, Tigist Zerihun⁶, Charlotte Hanlon^{4,7,8,9} and Rosa A. Hoekstra^{1,9} 

¹Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; ²Department of Biostatistics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; ³Centre for Implementation Science, Health Services and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; ⁴Department of Psychiatry, WHO Collaborating Centre for Mental Health Research and Capacity-Building, School of Medicine, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia; ⁵Department of Psychiatry, Yekatit 12 Hospital and Medical College, Addis Ababa, Ethiopia; ⁶Department of Psychiatry, St. Paul's Hospital Millennium Medical College, Addis Ababa, Ethiopia; ⁷Centre for Innovative Drug Development and Therapeutic Trials for Africa (CDT-Africa), College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia; ⁸Centre for Global Mental Health, Department of Health Services and Population Research and WHO Collaborating Centre for Mental Health Research and Capacity-Building, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK and ⁹King's Global Health Institute, King's College London, London, UK

Abstract

Background. Neurodevelopmental disorders (NDDs) are conditions affecting a child's cognitive, behavioural, and emotional development. Appropriate and validated outcome measures for use in children with NDDs in sub-Saharan Africa are scarce. The aim of this study was to validate the Communication Profile Adapted (CP-A), a measure developed in East Africa to assess caregivers' perception of communication among children with NDDs.

Methods. We adapted the CP-A for use in Ethiopia, focusing on the communicative mode (CP-A-mode) and function (CP-A-function) scales. The CP-A was administered to a representative sample of caregivers of children with NDDs and clinical controls. We performed an exploratory factor analysis and determined the internal consistency, test-retest reliability, within-scale, known-group, and convergent validity of the identified factors.

Results. Our analysis included $N = 300$ participants ($N = 139$ cases, $N = 139$ clinical controls, $N = 22$ who did not meet criteria for either cases or controls). Within the CP-A-mode, we identified two factors (i.e. *verbal* and *physical communication*); the CP-A-function scale was unidimensional. Combining both scales into one summary variable (the CP-A-total) resulted in a scale with excellent internal consistency and test-retest reliability (Cronbach's $\alpha = 0.97$; Kappa = 0.60–0.95, $p < 0.001$; ICC = 0.97, $p < 0.001$). Testing known-group validity, the CP-A-total scores were significantly higher for controls than cases (Δ mean = 33.93, $p < 0.001$). Convergent validity assessment indicated that scores were negatively and moderately correlated with clinical severity ($\rho = -0.25$, $p = 0.04$).

Conclusion. The CP-A is a valid tool for the assessment of communication among children with NDDs in Ethiopia. It holds promise as a brief, quantitative, and culturally appropriate outcome measure for use in sub-Saharan Africa.

Background

Neurodevelopmental disorders (NDDs) are a group of conditions that affect a child's cognitive, behavioural, and emotional development which include intellectual disability (ID), autism spectrum disorder (ASD), and attention deficit hyperactivity disorder (ADHD) (World Health Organization, 2019). According to Global Burden of Disease (GBD) 2016 estimates, developmental disabilities, including NDDs, affect 52.9 million children under 5 years of age. Of these, 95% live in low and middle-income countries (LMICs). Nearly 15 million of these children live in sub-Saharan Africa (Olusanya *et al.*, 2018). Ethiopia is one of the top 10 nations globally, with an estimated 1.3 million children living with developmental disabilities (Olusanya *et al.*, 2018). Children affected by NDDs and their caregivers experience severe stigma and social exclusion; this is especially true for those with ID and/or ASD (Tekola *et al.*, 2016, 2020; Tilahun *et al.*, 2016).

Despite the high prevalence and burden of NDDs, there is a wide gap between needs and service provision in sub-Saharan Africa. This applies to both fund allocation and the

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

availability of trained personnel (World Health Organization, 2013; Strand *et al.*, 2016; Chisholm *et al.*, 2019). Ethiopia, with a population of nearly 110 million, is served by only one specialist child psychiatrist. There are 0.08 general psychiatrists and 0.04 psychologists per 100000 people, but these cadres of workers have no specialist expertise in NDDs (World Health Organization, 2018).

This gap in resources and services for children with NDDs extends to research (Patel *et al.*, 2013; Tomlinson *et al.*, 2014). Only a negligible fraction of research on child development and mental health is conducted in LMICs (Kieling *et al.*, 2011; Durkin *et al.*, 2015; Nielsen *et al.*, 2017; Hoekstra *et al.*, 2018). This research gap results in an incomplete and biased body of knowledge (Durkin *et al.*, 2015; Hoekstra *et al.*, 2018). When evidence-based tools are lacking, diagnosis and intervention initiation tend to occur later or not occur at all, potentially impairing the prognosis and increasing the risk of comorbidities in affected children (Ruparelia *et al.*, 2016; Guralnick, 2017).

The development and evaluation of contextually appropriate interventions and harmonised and contextually valid outcome measures are essential to effectively address the service gap (Kieling *et al.*, 2011; Tomlinson *et al.*, 2014). There is currently no consensus on which outcome measures should be used in the evaluation of interventions targeting NDDs (Kohli-Lynch, Tann and Ellis, 2019). There is an urgent need for tools that are both accessible and appropriate for use in low-resource settings. Such tools are recommended to be: (i) of high quality, (ii) open-source and open-access, (iii) culturally appropriate, (iv) intuitive, (v) brief, (vi) acceptable, and (vii) easy to administer (Prince, 2013; Durkin *et al.*, 2015; Ruparelia *et al.*, 2016; de Vries, 2016; Carruthers *et al.*, 2018; Bakolis *et al.*, 2019; Kohli-Lynch *et al.*, 2019). These criteria rule out many of the existing outcome measures, which are often prohibitively expensive or rely on administration by highly qualified specialists (Durkin *et al.*, 2015).

In this study, we validate two scales of the Communication Profile-Adapted (CP-A) as a brief, culturally appropriate, caregiver-reported outcome measure for Ethiopia. The CP-A assesses caregivers' perceptions of their child's abilities and activities for communication, and participation in family and community events (Bunning *et al.*, 2014). It was developed to assess communication in children with complex communication needs and is thus suitable for use in children with NDDs. It has yet to be validated and assessed in its psychometric properties. The CP-A meets important criteria for use in low-resource settings. It has a solid theoretical background, based on the International Classification of Functioning, Disability and Health Framework (World Health Organization, 2001; Hartley and Wirz, 2002; Bunning *et al.*, 2014). It also meets the requirement of cultural relevance for sub-Saharan Africa as it was developed in Uganda and Kenya (Baker and Hartley, 1998, 1999; Bunning *et al.*, 2014). Because it is an open-access tool, translations and adaptations can be readily made. It is also easy to administer and does not use technical terminology (Bunning *et al.*, 2014).

We selected those scales of the CP-A that most closely reflect the key aspects to be targeted in interventional investigations for children with NDDs and hold promise as quantitative scales. These scales focus on the child's communicative mode and function. Both were adapted to the local context and assessed for validity and reliability. We hypothesised that children with NDDs would score significantly lower on the CP-A than controls and that CP-A scores would be negatively correlated with the severity of clinically diagnosed NDDs.

Methods

Setting

This study was carried out in Addis Ababa, at Yekatit 12 and St Pauls Millennium Medical College government hospitals between August 2018 and May 2019. The ethics protocol was approved by the College of Health Sciences Institutional Review Board at Addis Ababa University (062/16/Psy) and King's College London (HR-16/17-3489). In addition to validation of the CP-A, the data collection comprised further questionnaires, the validation of which are reported in Borissov *et al.* (2021).

Participants

Participants were 300 caregivers with long-term responsibilities for children aged 2–9 years, attending either the general paediatric or child mental health clinic of the two hospitals. The paediatric clinics consecutively recruited children with physical health conditions for the clinical control group. The mental health clinics consecutively recruited children with either NDDs alone or with a comorbid mental health condition for the case group, in accordance with the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (American Psychiatric Association, 2013). Given that no standardised diagnostic tests are available in Ethiopia, health practitioners relied on their clinical judgment based on interviews with caregivers, and observations and interactions with children. The clinical diagnoses in the mental health clinic were provided by general psychiatrists without specialist expertise in child psychiatry, as this specialty training is not available in Ethiopia.

Eligible families were approached by the attending clinician at either clinic and given a flyer with information about the study. Data collection and consent taking was done by clinic nurses who worked independently from the clinicians supporting the families. A subgroup of 40 caregivers was invited for a retest of an average of 19.6 days (s.d. = 3.8) after initial test data collection. All participants provided written informed consent.

Measures

CP-A

The CP-A is composed of 51 questions, divided into 3 main sections named 'body function and structure', 'activities for communication', and 'participation'. The portion 'activities for communication' includes six scales. Two of those, namely communicative mode (CP-A-mode) and communicative function (CP-A-function) were selected for validation in this study, see Tables 2 and 3 for item content. We refer to the combined items of these two scales as the CP-A-total.

Items' responses refer to a 0–4 Likert scale where [0] stands for 'never', [1] 'rarely', [2] 'sometimes', [3] 'usually', and [4] 'always'. Participants were shown a visual ladder representation of this scale with 0 being the lowest and 4 the highest rung (online Supplementary material Fig. S1).

Demographic information

Clinician-assigned severity levels were available for a subgroup of participants with a diagnosis of ID ($N = 24$) and ASD ($N = 50$). General psychiatrists working in the two hospitals rated the child's condition, based on the DSM-5 severity ratings, as [1] 'mild', [2] 'moderate', or [3] 'severe'. A structured questionnaire was used to collect caregiver-reported demographic information.

Procedure

CP-A adaptation, translation, and pre-testing

The CP-A was translated to Amharic, one of Ethiopia's official languages, and adapted to the local context. The instrument was considered by a consensus committee comprising native Amharic speakers fluent in English with expertise in the field. Following forward and backward translation, the committee met to formulate a version for pre-testing. This draft was pre-tested with 20 participants through cognitive interviews to understand how respondents interpreted instructions and items (Willis, 1999). Feedback was recorded by data collectors and subsequently discussed in the committee to establish a final draft.

Within the CP-A-mode, the content of question 1j was changed to 'Amharic' to reflect the main language spoken in Addis Ababa. Item 1, *Behaviour*, was followed by examples of challenging behaviours in brackets (e.g. crying, shouting) as cognitive interviews indicated that the direct translation of the term 'behaviour' was unclear to respondents. The original CP-A response scale comprised 8 rungs of ladders; this was shortened to 4 rungs (online Supplementary material Fig. S1) after cognitive interviewing indicated that participants could not meaningfully differentiate between the small increments of the original ladder. Within the CP-A-function, the original version of the scale did not just require a response using the ladder but also asked the extent to which the item is a problem for the respondent. This second part was removed to improve consistency as well as administration ease and time. In its original format, the CP-A asked both questions on how the child communicates with the respondent, and how the respondent communicates with the child. Caregivers participating in the cognitive interviews had difficulty distinguishing between these questions. In response, the section on how the caregiver communicates with the child was removed, focusing solely on the mode of communication of the child.

Data collection and entry

Data collection took place in Yekatit 12 and St Paul's Millennium Medical College general paediatric and child mental health clinics. To allow for the participation of non-literate caregivers, instruments were administered through face-to-face interviews by nurses who had received training in data collection procedures and questionnaire administration. Participating caregivers were reimbursed for their travel costs. Data were double entered using Epidata, version 3.1 (Christiansen and Lauritsen, 2003), to reduce the risk of data errors.

Data analyses

Demographics and item checks

The data were analysed using STATA, version 16 (Stata Corp, 2019). Group differences in demographic variables between cases and controls were explored using unpaired *t* test, Mann-Whitney test, Chi-Square, or Fisher's exact test depending on variables' characteristics. Missing values ($N=3$) in the CP-A were replaced by median imputation (Zhang, 2016).

Exploratory Factor Analysis (EFA)

The factor structure of the (i) CP-A-mode, (ii) CP-A-function, and (iii) CP-A-total were examined with the use of item exploratory factor analysis (EFA) (Bartholomew *et al.*, 2008). For each, we evaluated the polychoric matrix of their respective items. Exploratory factor analysis was applied to the matrix of item

correlation coefficients, to identify possible underlying dimensions. We used two criteria to aid the choice of the number of factors and provide empirical support for the selection: the scree plot and the criterion of eigenvalues above 1. χ^2 was used as a goodness-of-fit test to evaluate the adequacy of the number of extracted factors (Pett *et al.*, 2003). To assist data interpretation promax oblique rotation was used (Bartholomew *et al.*, 2008; Baldwin, 2019). Factors were labelled referring to theoretical notions and interpretability.

Validity

Our approach of studying the validity and reliability was guided by the consensus-based standards for the selection of health measurement instruments (COSMIN) guidelines (Mokkink *et al.*, 2010, 2019).

Within-scale validity. We assessed within-scale validity (the extent to which the subscales of an instrument measure the same concept; Brohan *et al.*, 2013)) by examining the correlation between summative scores of the identified factors as well as those of the CP-A-mode and CP-A-function.

Known-group validity. Known-group validity (i.e. the ability to distinguish among distinct groups; Streiner *et al.*, 2015) was tested by assessing group differences between the NDDs and clinical control group across identified factors. Unadjusted and adjusted mean differences between cases and controls were tested using multivariable linear regression, with group membership and previously identified demographic characteristics as independent variables.

Convergent validity. Convergent validity (how related different measures assessing associated constructs are; Streiner *et al.*, 2015), was assessed by estimating the correlation between the total scores of the identified factors and clinician-rated clinical severity measures available for a subset of children with ID or ASD. The clinical severity scores for each disorder were merged into a single score to maximise sample size; in the case of double diagnoses, the highest severity score was retained.

Reliability

Internal consistency (the extent to which items of the same scale measure the same construct) was evaluated through Cronbach's alpha (α) (Revicki, 2014). Test-retest reliability (the agreement between scores on the same scale across timepoints) was assessed through weighted Kappa and Interclass Correlation Coefficient (ICC) (Streiner, Norman and Cairney, 2015). The ICC was calculated for continuous summary scores, weighted Kappa for single categorical items.

Results

Demographics and item checks

Our sample included $N=300$ caregivers, comprising $N=139$ cases and $N=139$ clinical controls. Children in the case group presented with ASD, ID, ADHD, language delay, global developmental delay, and/or Down's syndrome, see Table 1 (severity ratings details are provided in online Supplementary material Table S1). Those in the clinical control group were affected by a range of physical health problems (e.g. cardiovascular, respiratory, neurological conditions; details in online Supplementary

Table 1. Conditions reported in the case group including comorbidities.

Neurodevelopmental disorders (N=139)		
Condition	N	%
ASD	55	40
ADHD	28	20
ID	19	17
LD	6	4
DS	1	0.7
GDD	1	0.7
ASD + ID	11	8
ASD + ADHD	6	4
ASD + LD	2	1
ASD + GDD	1	0.7
ADHD + ID	2	1
ADHD + LD	1	0.7
ID + DS	2	1
ID + GDD	1	0.7
DS + GDD	1	0.7
ASD + ADHD + ID	1	0.7
ADHD + ID + DS	1	0.7
Total	139	100

Note. ASD, autism spectrum disorder; ADHD, attention deficit hyperactivity disorder; ID, intellectual disability; LD, language disorder; DS, Down's syndrome; GDD, global developmental delay. N, number of observations; %, frequency

Table S2). The remaining participants (rest: $N = 22$) included children who did not meet the criteria to be in either the case or clinical control group (e.g. they were diagnosed with epilepsy or a mental health condition, but did not have an NDD) and those for whom diagnostic information was incomplete (full details in online Supplementary Table S3).

Table 2 presents the demographic characteristics of the caregivers and children. The majority of caregivers were females, mainly mothers of the child, and housewives. Most caregivers were married, living in urban areas, and of Orthodox Christian religion. Most did not receive any education above the primary school level. The mean age of caregivers at the time of the interview was of 34 years ($S.D. = 7.1$).

Statistically significant differences between cases and controls were observed for the caregivers' gender ($p < 0.001$), with a higher proportion of females among cases, and residence ($p < 0.001$), with more cases living in urban areas. Furthermore, we report significant differences in religion, ($p = 0.004$), observing a higher proportion of Muslims among cases compared to controls, relationship to the child ($p < 0.001$), where caregivers in the case group were more often mothers, and occupation ($p = 0.02$), with a higher proportion of housewives among cases.

62% of children were male, with an average age of 4.9 years ($S.D. = 2.0$). Among children, comparison across cases and controls showed significant differences in age ($p < 0.001$), with cases being older than controls, and gender ($p < 0.001$), with a higher proportion of males in the case group in line with the

observation that NDDs are more common in boys (Loomes *et al.*, 2017; Sayal *et al.*, 2018). The retest sample consisted of $N = 40$ caregivers ($N = 19$ cases, $N = 20$ controls, $N = 1$ rest).

Items *i* 'Sign language', *j* 'Speaking English', and *m* 'Other' from the CP-A-mode, had a median and iqr of 0 across test and retest, suggesting no or limited variability. These items were therefore removed. The remaining $N = 10$ items for CP-A-mode and $N = 23$ items for CP-A-function were further analysed.

Exploratory factor analysis

The Scree-test and eigenvalues (online Supplementary material Fig. S2) suggested a 2-factor solution for the CP-A-mode ($\chi^2 = 1501.84$, $p < 0.001$). The correlations between items and promax-rotated common factors are displayed in Table 3. Factor loadings and structure matrix indicated that items *a* 'Facial expression', *c* 'Gestures', *d* 'Body movements', *e* 'Looking or use of eye gaze', *f* 'Pointing', and *h* 'Showing you objects' loaded more strongly on factor 1. Items *b* 'Making noises (vocals)', *g* 'Showing you pictures', *j* 'Speaking Amharic' and *l* 'Behaviour', instead loaded on factor 2. Factor 1 generally corresponds to *physical communication*, while factor 2 to *verbal communication*. Item *g* 'Showing you pictures' presented moderate loadings for both factors (slightly higher for the *verbal communication*) that were maintained throughout different rotation methods. Item *l* had a negative loading on *verbal communication*, indicating that caregivers of children with poor verbal communication tended to highlight the use of behaviour as the main mode of their child's communication.

For communicative function, the Scree-test and eigenvalues (online Supplementary material Fig. S3) suggested a 1-factor solution. The model had a good fit to the data ($\chi^2 = 8784.87$; $p < 0.001$), suggesting that the function items assessed a single construct. All items loaded positively on the factor, see Table 3.

Similarly, for the CP-A-total the Scree-test and eigenvalues (online Supplementary material Fig. S4) indicated a 1-factor solution. The model had a good fit to the data ($\chi^2 = 2.1e + 04$; $p < 0.001$), suggesting that all items of the analysed CP-A scales can be meaningfully subsumed under a single construct reflecting overall communication activities. All items, besides Item *l*, loaded positively on the factor (online Supplementary material Table S4).

Validity

Within-scale validity

A summary of the psychometric properties of the CP-A is provided in Table 4. The Spearman's rank correlation coefficient (ρ) between the CP-A-mode's *verbal communication* and *physical communication* scores was significant and moderately positive ($\rho = 0.41$, $p < 0.001$). Similarly, the correlation between the CP-A-mode and CP-A-function scores was significant and moderately positive ($\rho = 0.59$, $p < 0.001$), further justifying the adoption of the CP-A-total as a unidimensional summary scale (Lamping *et al.*, 2002; Dancey and Reidy, 2017).

Known-group validity

The multiple regression analyses indicate that clinical controls scored significantly higher than cases on CP-A-mode's *physical communication* ($\Delta\text{mean} = 2.89$, $p < 0.001$) and *verbal communication* ($\Delta\text{mean} = 2.09$, $p < 0.001$). This holds true also for the CP-A-function ($\Delta\text{mean} = 28.94$, $p < 0.001$) and CP-A-total ($\Delta\text{mean} = 33.93$, $p < 0.001$), Table 5. The significant group differences persisted when adjusting for covariates (caregiver's age,

Table 2. Demographic information

	Control		Case		Rest*	
	139		139		22	
Total	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Caregiver						
Gender***						
Male	43	30.9	14	10.1	7	31.8
Female	96	69.1	125	89.9	15	68.2
Marital Status						
Married	122	87.8	111	79.9	18	81.8
Single	4	2.9	5	3.6	1	4.6
Divorced	10	7.2	18	13.0	3	13.6
Widowed	3	2.2	4	2.9	0	0
Missing	0	0	1	0.7	0	0
Residence***						
Rural	45	32.4	18	13.0	6	27.3
Urban	91	65.5	118	84.9	16	72.7
Missing	3	2.2	3	2.2	0	0
Religion**						
Orthodox Christian	95	68.4	79	56.8	15	68.2
Protestant	18	13.0	11	7.9	1	4.6
Catholic	1	0.7	1	0.7	0	0
Muslim	23	16.6	47	33.8	6	27.3
Other (Please Specify)	2	1.4	0	0	0	0
Missing	0	0	1	0.7	0	0
Relationship to child***						
Mother	94	67.6	116	83.5	14	63.6
Father	41	29.5	13	9.4	6	27.3
Extended family	4	2.9	5	3.6	2	9.1
Other	0	0	3	2.2	0	0
Missing	0	0	2	1.4	0	0
Education						
No formal education	20	14.4	20	14.4	4	18.2
Primary school (Grade 1-8)	49	35.3	42	30.2	12	54.6
Secondary school (Up To Grade 10 + 2 Or 12 + 1)	46	33.1	47	33.8	4	18.2
Diploma (Grade 10 + 3 And Above Or 12 + 2 And Above)	12	8.6	11	7.9	1	4.6
College	9	6.5	17	12.2	0	0
Missing	3	2.2	2	1.4	1	4.6
Occupation*						
Farmer	7	5	5	3.6	4	18.2
Housewife	41	29.5	72	51.8	13	59.1
Merchant	15	10.8	16	11.5	2	9.1
Student	1	0.7	0	0	0	0
Civil servant	11	7.9	6	4.3	1	4.6

(Continued)

Table 2. (Continued.)

	Control		Case		Rest*	
	139		139		22	
Total	N	%	N	%	N	%
Daily labourer	19	13.7	11	7.9	0	0
Other	8	5.8	4	2.9	0	0
Missing	37	26.6	25	18	2	9.1
	Mean (s.d.)	Range	Mean (s.d.)	Range	Mean (s.d.)	Range
Age (years)	33.4 (6.7)	17–63	34.9 (7.6)	18–62	34.6 (6.8)	22–50
Number of children under care	2.4 (1.2)	1–8	2.4 (1.2)	1–10	3.2 (1.8)	1–8
Child						
Gender***						
Male	69	49.6	104	74.8	13	59.1
Female	69	49.6	35	25.2	9	40.9
Missing	1	0.7	0	0	0	0
	Mean (s.d.)	Range	Mean (s.d.)	Range	Mean (s.d.)	Range
Age (years)***	4.3 (2.0)	2–9	5.3 (1.8)	2–9	5.6 (2.0)	2–9

Note: *Rest group refers to 22 participants included in the EFA analyses that did not meet inclusion criteria for either the case or control group. Significant differences between cases and controls are displayed as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; N, number of observations; %, frequency; s.d., standard deviation.

religion, occupation, residence, and relationship to the child; children's age and gender). Cohen's d estimates for CP-A-mode's factors were medium when unadjusted and small when accounting for covariates. Estimates were large for CP-A-function and CP-A-total across both scenarios (Cohen, 1988).

Convergent validity

There was a significant correlation between clinical severity and CP-A-mode's *physical communication* ($\rho = -0.24$, $p = 0.04$) as well as *verbal communication* ($\rho = -0.27$, $p = 0.02$). Similar results were obtained for the CP-A-total ($\rho = -0.25$, $p = 0.04$). For the CP-A-function, the point estimate of the correlation was in the same direction and of similar magnitude ($\rho = -0.23$) but was not significant ($p = 0.59$) (Dancey and Reidy, 2017).

Reliability

The internal consistency for the CP-A-mode factors was acceptable (*physical communication*: 4 items, $\alpha = 0.80$, 95% confidence interval (CI) ≥ 0.78 ; *verbal communication*: 6 items, $\alpha = 0.72$, 95% CI ≥ 0.69). Excellent internal consistencies were obtained for the CP-A-function (23 items, $\alpha = 0.96$, 95% CI ≥ 0.96) and the CP-A-total (33 items, $\alpha = 0.97$, 95% CI ≥ 0.96). Assessing internal consistency separately for cases and controls suggested that levels remained acceptable for the CP-A-mode's *physical communication* (cases: $\alpha = 0.78$, 95% CI ≥ 0.74 ; controls: $\alpha = 0.78$, 95% CI ≥ 0.74), but were low for *verbal communication* (cases: $\alpha = 0.63$, 95% CI ≥ 0.57 ; controls: $\alpha = 0.64$, 95% CI ≥ 0.57). The internal consistency remained excellent for CP-A-function (cases: $\alpha = 0.94$, 95% CI ≥ 0.93 ; controls: $\alpha = 0.96$, 95% CI ≥ 0.95) and CP-A-total (cases: $\alpha = 0.94$, 95% CI ≥ 0.95 ; controls: $\alpha = 0.94$, 95% CI ≥ 0.93) (George and Mallery, 2010; Revicki, 2014).

Weighted Kappa coefficients ranged between 0.60 and 0.86 for *physical communication* and between 0.65 and 0.76 for *verbal communication* ($p < 0.001$), demonstrating moderate to a substantial agreement among items. For CP-A-function and CP-A-total, the agreement was moderate to near perfect for all items (min = 0.60, max = 0.95; $p < 0.001$) (McHugh, 2012; Portney, 2020). Please refer to online Supplementary material Tables S5, S6 for details.

The ICC indicated good test-retest reliability for the CP-A-mode's *physical communication* (ICC = 0.81, 95% CI 0.67–0.89; $p < 0.001$) and *verbal communication* (ICC = 0.83; 95% CI 0.70–0.91; $p < 0.001$). Excellent test-retest reliability was observed for the CP-A-function (ICC = 0.96; 95% CI 0.94–0.98; $p < 0.001$) and CP-A-total (ICC = 0.97; 95% CI 0.95–0.98; $p < 0.001$) (Koo and Li, 2016; Portney, 2020).

Discussion

This paper reports the first validation study of the CP-A, a caregiver-reported tool for the assessment of children's communication. Our aim was to address the need for a psychometrically sound, brief, and culturally appropriate outcome measure for use in Ethiopia. We investigated two sections of the CP-A: communicative mode and function. Within communicative mode we identified two factors, *verbal* and *physical communication*; the communicative function scale was unidimensional. EFA results indicated that all items, from both scales, can be meaningfully summarised into one single factor. This suggested the adoption of a summary score (CP-A-total), supported by findings of strong correlations between the identified factors. As hypothesised, children with NDDs (cases) scored lower than clinical controls. Moreover, scale scores were negatively correlated to clinical severity ratings of NDDs, indicating that children with more severe

Table 3. Factor loadings of the items of the CP-A-mode and CP-A-function

Communicative Mode		
Item	Factor 1 <i>Physical communication</i>	Factor 2 <i>Verbal communication</i>
<i>1. How does (child's name) communicate with you?</i>		
a. Facial expression	0.55	0.53
b. Making noises (<i>vocals</i>)	0.33	0.61
c. Gestures	0.73	0.37
d. Body movements	0.65	0.40
e. Looking or use of eye gaze	0.67	0.58
f. Pointing	0.79	0.27
g. Showing you pictures	0.50	0.59
h. Showing you objects	0.72	0.22
j. Speaking <i>Amharic, Guragigna meskan and Mareko</i>	0.40	0.89
l. Behaviour	-0.17	-0.71
Communicative function		
Item	Factor 1 <i>Communicative function</i>	
<i>Expressive</i>	1. Does (child's name) let you know if s/he doesn't like something you are giving him/her?	
	0.78	
	2. Does (child's name) let you know when s/he is sad or upset?	
	0.73	
	3. Does (child's name) indicate 'yes', for example, if s/he wants to do something?	
	0.88	
	4. Does (child's name) indicate 'no', for example, if s/he doesn't want to do something?	
	0.83	
	5. Does (child's name) let you know if s/he is happy?	
	0.81	
	6. Does (child's name) let you know if s/he is not happy?	
	0.77	
	7. Does (child's name) get your attention when s/he wants?	
	0.50	
	8. Does (child's name) tell you what s/he wants, for example, food or drink?	
	0.80	
	9. Does (child's name) greet people?	
	0.80	
	10. Does (child's name) get you to do something again which s/he has just enjoyed doing?	
	0.70	
	11. Does (child's name) ask for help when s/he can't manage to do something by him/herself?	
	0.61	
	12. Does (child's name) comment on things that are happening	
	0.87	
	13. Does (child's name) ask simple questions, for example, what is it or where's mummy?	
	0.88	
	14. Does (child's name) tell you about something that has happened, for example, when you weren't looking?	
	0.90	
<i>Social</i>	1. Does (child's name) start up communication with people in the family?	
	0.87	
	2. Does (child's name) start up communication with other people?	
	0.88	
	3. Does (child's name) try again if you don't understand him/her?	
	0.82	
	4. Does (child's name) communicate with other people in a way that is polite?	
	0.89	
<i>Receptive</i>	1. Does (child's name) understand when you tell him/her 'no'?	
	0.79	
	2. Does (child's name) understand simple instructions?	
	0.87	
	3. Does (child's name) understand if you ask for something that is not in the immediate environment?	
	0.87	
	4. Does (child's name) understand if you communicate about something that is going to happen or has already happened?	
	0.86	
	5. Does (child's name) understand stories?	
	0.88	

Note: Factor loadings extracted through principal axis factoring. Shaded cells indicate factor allocation.

Table 4. Psychometric properties of the CP-A's identified factors

Psychometric property	Constituent parts	CP-A-mode's physical communication	CP-A-mode's verbal communication	CP-A-function	CP-A -total
		N = 300			
Validity	Known-group validity (Δ mean)	2.89**** ^a	2.09**** ^a	28.94**** ^a	33.93**** ^a
	Convergent validity (ρ)	-0.24*	-0.27*	-0.23 NS	-0.25*
Reliability	Internal Consistency (Cronbach's alpha)	0.80 ^b	0.72 ^c	0.96 ^d	0.96 ^e
	Test-retest reliability (Weighted Kappa)	0.60-0.86****	0.65-0.76****	0.60-0.95****	0.60-0.95****
	Test-retest reliability (Interclass Correlation Coefficient)	0.81****	0.83****	0.96****	0.97****

Note: Significant differences between cases and controls are displayed as * $p < 0.05$, ** $p < 0.01$, **** $p < 0.001$; N, number of participants.

^aUnadjusted mean difference in summative scores (controls minus cases).

^b6 items.

^c4 items.

^d23 items.

^e33 items.

NDDs used fewer modes of communication and applied fewer functions of communication. We observed acceptable to excellent internal consistency as well as test-retest reliability. Overall, these results demonstrated the validity and reliability of the CP-A-mode, function, and communication-total scales.

Within the CP-A-mode, for items *i* 'Sign language', *j* 'Speaking English', and *m* 'Other', most responses were equal to 0 (i.e. 'never'), suggesting these items have little relevance in assessing communicative mode within our sample. For item *i* this is likely due to the fact that none of our participants was reported to have hearing loss. Furthermore, formal sign language has received very limited implementation in Ethiopia so far (Wakuma, 2015), and caregivers could have selected item *c* 'Gestures' to indicate informal signs as modes of communication. The lack of variability in item *j* can be attributed to the fact that English is not widely spoken across the population (Central Intelligence Agency, 2020). Most caregivers answered 'never' to item *m* 'Other', suggesting that all previously administered questions had exhaustively described the communicative modes adopted by their children. For these reasons, items *i*, *j*, and *m* of the communicative mode were dropped from subsequent analyses. The CP-A-function scale was retained in its entirety.

For communicative mode, we found support for a 2-factor structure. All factor 1 items fit with the profile for *physical communication*. For factor 2, the construct of *verbal communication* is defined by the items with the highest loadings, *b* 'Making noises (vocals)' and *j* 'Speaking Amharic'. Item *g* 'Showing you pictures' showed more moderate loading, with substantial cross-loading on factor 1. Item *l* 'Behaviour' had a strong negative correlation with factor 2, suggesting that caregivers of children who did not express themselves verbally were more likely to report their child's behaviour as a form of communication. This finding is in line with theoretical notions that see verbal acts and behaviours as equivalent in function (Carr and Durand, 1985). When verbal communication is severely impaired, behavioural expression may become challenging to the person and others (Royal College of Psychiatrists and Banks, 2007). The occurrence of challenging behaviours (e.g. self-injury, stereotypy) is reported across NDDs and cultural contexts (McClintock et al., 2003; Adeniyi and Omigbodun, 2016; O'Nions et al., 2018). The Amharic version of the item, contrary to the English original, was followed by examples of challenging behaviours in brackets '(e.g. crying and

shouting)'. These examples may have contributed to caregivers' interpretation of the item as primarily concerning challenging behaviours rather than behaviour overall.

Clinical controls scored higher than cases on the CP-A-mode, function, and communication-total scales, even after adjusting for covariates. This supports both our initial hypothesis and that of the developers of the measure: these scales were designed to reflect higher perceived competence in communication through higher summative scores (Bunning et al., 2014). This is in line with studies investigating other tools assessing communication in higher-income countries (HICs), with lower ratings consistently indicating more profound impairments (Geurts et al., 2004; Norbury et al., 2004).

Significance

This investigation represents the first exploration of the validity and psychometric properties of the CP-A. Compared to other measures, it is more suitable for application in low-resource settings. Unlike caregiver-reported tools developed in Western HICs that assess similar constructs, the CP-A is free and open access, and this avoids the significant costs and adaptation negotiations associated with copyright-restricted instruments (Durkin et al., 2015). Moreover, instruments developed in Western HICs often require extensive adaptations to be suitable in non-Western lower-income contexts (Marlow et al., 2019; de Leeuw et al., 2020). As the CP-A is one of the very rare measures developed in an LMIC (Bunning et al., 2014; Goldfeld and Yousafzai, 2018), it does not encounter these issues. Its design and content are more likely to be relevant and appropriate for the African context. Nevertheless, adaptations of limited entities (e.g. referring to the languages spoken locally) are required to fit the specific context of an application. These must be paired with further explorations of the psychometric properties in diverse settings across the continent.

The only previously published research using the CP-A as an outcome measure investigated the impact of a caregiver-driven intervention for children with complex communication needs in rural Kenya (Bunning et al., 2014). While the sample size of this study was small ($N = 10$) and did not include a control group, results suggested sensitivity to change, as scores in the activities for communication sections, including communicative mode and function, were significantly higher post-intervention

Table 5. Adjusted and unadjusted mean differences in summative scores between cases and controls

	N	mean (s.d.) cases	mean (s.d.) controls	Unadjusted			Adjusted		
				Δ mean	95% CI	Cohen's d	Δ mean	95% CI	Cohen's d
CP-A-mode's physical communication	6	11.21 (5.84)	14.10 (4.04)	2.89***	(1.71, 4.08)	0.58	2.25**	(0.62, 3.89)	0.33
CP-A-mode's verbal communication	4	6.85 (3.18)	8.94 (2.16)	2.09***	(1.45, 2.73)	0.77	1.61***	(0.69, 2.52)	0.42
CP-A-function	23	36.67 (19.81)	65.61 (17.31)	28.94***	(24.55, 33.34)	1.56	30.42***	(24.60, 36.23)	1.24
CP-A-total	33	54.73 (25.82)	88.65 (19.77)	33.93***	(28.49, 39.36)	1.48	34.28***	(26.93, 41.62)	1.10

Note: Significant differences between cases and controls are displayed as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Δ mean, mean difference (controls minus cases); N, number of items. Mean differences unadjusted or adjusted for covariates.

compared to pre-intervention (Bunning *et al.*, 2014). The integration of their findings with that of our investigation demonstrates the potential of the CP-A for use in interventional studies. Adopting an accessible and appropriate measure like the CP-A across investigations on NDDs would increase the comparability of results, aiding the evaluation and implementation of effective interventions in low-resource settings.

Limitations

Limitations should be considered when evaluating our results. Significant differences were reported for some demographic variables, especially in terms of age, where our clinical control group was younger than the case group. Nevertheless, it could be argued that the developmental age of the two groups is more comparable in this situation, given that the control group (with younger children that are naturally in the earlier phases of developing their communication abilities) still scores higher than the case group. The clinical severity assessment was carried out by general psychiatrists, with no specialist expertise in the assessment of child NDDs. Nevertheless, diagnosis and assessment of children with NDDs part of postgraduate training in psychiatry in Ethiopia and the psychiatrists involved in the study were experienced in making these diagnostic assessments. Psychiatrists used their clinical judgement rather than standardised tools to assess the severity of impairment of the children since there are no validated standardised clinical severity assessment scales available in Ethiopia to support the assessment of severity. Furthermore, the limited educational or supportive service provision available for children with NDDs in this setting means that reports from other professionals are not available to inform severity assessments. Thus severity is based on a single report from the caregiver during the clinical encounter and observations of the psychiatrist of the child in a clinical setting. Severity scores were collected for a small sample only, covering a limited range of complex communication needs. This study was conducted as part of a larger project focusing on NDDs. Most cases included in our study had ASD or ID, rather than a wider group of complex communication needs and developmental disabilities for which the CP-A was also developed (e.g. sensory impairments). Lastly, participants were help-seeking families recruited in Addis Ababa. Our sample had an overrepresentation of urban families and may not be fully representative of the Ethiopian population.

Future research

Further research could test our factor structure for the CP-A-mode and CP-A-function through confirmatory factor analysis. Recruitment of cases should be extended to a wider range of complex communication needs. In such studies, the relevance of items reflecting characteristics of developmental disabilities with no representation in our sample (i.e. *i* 'Sign language'), should be re-evaluated. Future research should further examine whether the CP-A-mode, CP-A-function, and CP-A-total are sensitive to change induced by interventions. Lastly, future studies may also wish to consider other sections of the CP-A not included in the current evaluation.

Conclusion

This work is the first investigation to explore the validity of CP-A, an open-access measure developed in and for the African context.

The communication mode and function and their combined scales met the validity and reliability criteria as a measure for the assessment of caregiver-perceived activities for communication. We tested this among children with NDDs and concomitant complex communication needs in Ethiopia. We recommend the further validation of this scale. The CP-A has potential for application in intervention studies on NDDs across sub-Saharan Africa as a brief, quantitative, and culturally appropriate outcome measure.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/gmh.2021.44>

Acknowledgements. The authors are grateful to all caregivers for their participation.

Financial support. This work was supported by joint funding from the Medical Research Council (MRC) (United Kingdom), Department for International Development (DFID), Wellcome Trust and National Institute for Health Research (NIHR) (#MR/P020844/1). BT, CH and RAH receive support from the National Institute of Health Research (NIHR200842) and CH through the NIHR Global Health Research Unit on Health System Strengthening in Sub-Saharan Africa, King's College London (GHRU 16/136/54) using UK aid from the UK Government. The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. CH receives support from AMARI as part of the DELTAS Africa Initiative [DEL-15-01]. IB is supported by the NIHR Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust and by the NIHR Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust."

Conflict of interest. Conflicts of Interest: None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Adeniyi YC, and Omigbodun OO (2016) Effect of a classroom-based intervention on the social skills of pupils with intellectual disability in Southwest Nigeria. *Child and Adolescent Psychiatry and Mental Health*. [Online] **10**, 29.
- American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. [Online]. Washington, DC: American Psychiatric Association. doi: 10.1176/appi.books.9780890425596.
- Baker H, and Hartley S (1998) *Designing a communication ability outcome measuring tool* (Unpublished master's thesis). Centre of International Child Health, University of London.
- Baker H, and Hartley S (1999) Piloting a communication disability outcome measuring tool in Uganda. *Department for International Development (DFID), UK, Report, School-aged Child Health Programme*. London: Centre for International Child Health, Institute of Child Health.
- Bakolis I, Thornicroft G, Vitoratou S, Rüsçh N, Bonetto C, Lasalvia A, and Evans-Lacko S (2019) Development and validation of the DISCUS scale: a reliable short measure for assessing experienced discrimination in people with mental health problems on a global level. *Schizophrenia Research*. [Online] **212**, 213–220.
- Baldwin S (2019) *Psychological Statistics and Psychometrics Using Stata*, 1st Edn., College Station, Texas: Stata Press.
- Bartholomew DJ, Steele F, Moustaki I, and Galbraith JI (2008) *Analysis of Multivariate Social Science Data*, 2nd Edn., Boca Raton: Routledge.
- Borissov A, Bakolis I, Tekola B, Kinfe M, Ceccarelli C, Girma Bayouh F, Abdurahman R, Zerihun T, Hanlon C, and Hoekstra R (2021) Adaptation and validation of two autism-related measures of skills and quality of life in Ethiopia. *Autism*. (in press). doi: 10.1177/13623613211050751
- Brohan E, Clement S, Rose D, Sartorius N, Slade M and Thornicroft G (2013) Development and psychometric evaluation of the Discrimination and Stigma Scale (DISC). *Psychiatry Research*. [Online] **208**, 33–40.
- Bunning K, Gona JK, Newton CR, and Hartley S (2014) Caregiver perceptions of children who have Complex communication needs following a home-based intervention using augmentative and alternative communication in rural Kenya: an intervention note. *Augmentative and Alternative Communication*. [Online] **30**, 344–356.
- Carr EG, and Durand VM (1985) Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*. [Online] **18**, 111–126.
- Carruthers S, Kinnaird E, Rudra A, Smith P, Allison C, Auyeung B, Chakrabarti B, Wakabayashi A, Baron-Cohen S, Bakolis I, and Hoekstra RA (2018) A cross-cultural study of autistic traits across India, Japan and the UK. *Molecular Autism*. [Online] **9**, 52.
- Central Intelligence Agency (2020) *The World Factbook*. [Online]. Langley, VA: Central Intelligence Agency Publications. Available from: <https://www.cia.gov/the-world-factbook/countries/ethiopia/> [Accessed: 24 June 2020].
- Chisholm D, Docrat S, Abdulmalik J, Alem A, Gureje O, Gurung D, Hanlon C, Jordans MJD, Kangere S, Kigozi F, Mugisha J, Muke S, Olayiwola S, Shidhaye R, Thornicroft G, and Lund C (2019) Mental health financing challenges, opportunities and strategies in low- and middle-income countries: findings from the emerald project. *BJPsych Open*. [Online] **5**, e68.
- Christiansen TB, and Lauritsen JM (2003) *EpiData - Comprehensive Data Management and Basic Statistical Analysis System*. [Online]. Odense, Denmark: EpiData Association. Available at <https://www.epidata.dk/credit.htm> [Accessed: 26 November 2020].
- Cohen J (1988) Google-Books-ID: 2v9zDAsLvA0C. *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dancey C, and Reidy J (2017) Google-Books-ID: RSC9DgAAQBAJ. *Statistics Without Maths for Psychology*. Harlow, UK: Pearson Education Limited.
- de Leeuw A, Happe F, and Hoekstra RA (2020) A conceptual framework for understanding the cultural and contextual factors on autism across the globe. *Autism Research*. [Online] **13**, 473–476.
- de Vries PJ (2016) Thinking globally to meet local needs: autism spectrum disorders in Africa and other low-resource environments. *Current Opinion in Neurology*. [Online] **29**, 130–136.
- Durkin MS, Elsabbagh M, Barbaro J, Gladstone M, Happe F, Hoekstra RA, Lee L-C, Rattazzi A, Stapel-Wax J, Stone WL, Tager-Flusberg H, Thurm A, Tomlinson M, and Shih A (2015) Autism screening and diagnosis in low resource settings: challenges and opportunities to enhance research and services worldwide. *Autism Research*. [Online] **8**, 473–476.
- George D, and Mallery P (2010) Google-Books-ID: KS1DPgAACAAJ. *SPSS for Windows Step by Step: A Simple Guide and Reference*, 17.0 Update. Allyn & Bacon.
- Geurts HM, Verté S, Oosterlaan J, Roeyers H, Hartman CA, Mulder EJ, van Berckelaer-Onnes IA, and Sergeant JA (2004) Can the children's communication checklist differentiate between children with autism, children with ADHD, and normal controls?: cCC profiles in ADHD and autism. *Journal of Child Psychology and Psychiatry*. [Online] **45**, 1437–1453.
- Goldfeld S, and Yousafzai A (2018) Monitoring tools for child development: an opportunity for action. *The Lancet Global Health*. [Online] **6**, e232–e233.
- Guralnick MJ (2017) Early intervention for children with intellectual disabilities: an update. *Journal of Applied Research in Intellectual Disabilities*. [Online] **30**, 211–229.
- Hartley SD, and Wirz SL (2002) Development of a 'communication disability model' and its implication on service delivery in low-income countries. *Social Science & Medicine*. [Online] **54**, 1543–1557.
- Hoekstra RA, Girma F, Tekola B, and Yenus Z (2018) Nothing about us without us: the importance of local collaboration and engagement in the global study of autism. *BJPsych International*. [Online] **15**, 40–43.
- Kieling C, Baker-Henningham H, Belfer M, Conti G, Ertem I, Omigbodun O, Rohde LA, Srinath S, Ulkuer N, and Rahman A (2011) Child and adolescent mental health worldwide: evidence for action. *The Lancet*. [Online] **378**, 1515–1525.

- Kohli-Lynch M, Tann CJ, and Ellis ME** (2019) Early intervention for children at high risk of developmental disability in low- and middle-income countries: a narrative review. *International Journal of Environmental Research and Public Health*. [Online] **16**, 44–49. doi: 10.3390/ijerph16224449
- Koo TK, and Li MY** (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*. [Online] **15**, 155–163.
- Lamping DL, Schroter S, Marquis P, Marrel A, Duprat-Lomon I, and Sagnier P-P** (2002) The community-acquired pneumonia symptom questionnaire. *Chest*. [Online] **122**, 920–929.
- Loomes R, Hull L, and Mandy WPL** (2017) What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*. [Online] **56**, 466–474.
- Marlow M, Servili C, and Tomlinson M** (2019) A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low- and middle-income countries. *Autism Research*. [Online] **12**, 176–199.
- McClintock K, Hall S, and Oliver C** (2003) Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. *Journal of Intellectual Disability Research*. [Online] **47**, 405–416.
- McHugh ML** (2012) Interrater reliability: the kappa statistic. *Biochemia Medica*. [Online] **22**, 276–282. Available from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/> [Accessed: 12 June 2020].
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, and de Vet HCW** (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*. [Online] **19**, 539–549.
- Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, de Vet HC, and Terwee CB** (2019) COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 32.
- Nielsen M, Haun D, Kärtner J, and Legare CH** (2017) The persistent sampling bias in developmental psychology: a call to action. *Journal of Experimental Child Psychology*. [Online] **162**, 31–38.
- Norbury CF, Nash M, Baird G, and Bishop DVM** (2004) Using a parental checklist to identify diagnostic groups in children with communication impairment: a validation of the children's communication checklist – 2. *International Journal of Language & Communication Disorders*. [Online] **39**, 345–364.
- Olusanya B, Davis AC, Wertlieb D, Boo N-Y, Nair MKC, Halpern R, Kuper H, Breinbauer C, de Vries PJ, Gladstone M, Halfon N, Kancherla V, Mulaudzi MC, Kakooza-Mwesige A, Ogbo FA, Olusanya JO, Williams AN, Wright SM, Manguerra H, Smith A, Echko M, Ikeda C, Liu A, Milliar A, Ballesteros K, Nichols E, Erskine HE, Santomauro D, Rankin Z, Smith M, Whiteford HA, Olsen HE, and Kassebaum NJ** (2018) Developmental disabilities among children younger than 5 years in 195 countries and territories, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet Global Health*. [Online] **6**, e1100–e1121.
- O'Nions E, Happé F, Evers K, Boonen H, and Noens I** (2018) How do parents manage irritability, challenging behaviour, non-compliance and anxiety in children with autism spectrum disorders? A meta-synthesis. *Journal of Autism and Developmental Disorders*. [Online] **48**, 1272–1286.
- Patel V, Kieling C, Maulik PK, and Divan G** (2013) Improving access to care for children with mental disorders: a global perspective. *Archives of disease in childhood*. [Online] **98**, 323–327.
- Pett MA, Lackey NR, and Sullivan JJ** (2003) Google-Books-ID: 9kB5jE2jS4C. *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. Newbury Park, CA: SAGE.
- Portney LG** (2020) *Foundations of Clinical Research: Applications to Evidence-Based Practice*, 4th Edn., Philadelphia: F.A. Davis.
- Prince M** (2013) Cross-cultural research methods and practice. In Patel V, Prince M, Minas H and Cohen A (eds), *Global Mental Health: Principles and Practice*. Oxford, UK: Oxford University Press, pp. 63–81.
- Revicki D** (2014) Psychological statistics and psychometrics using stata. In Michalos AC (ed.), *Encyclopedia of Quality of Life and Well-Being Research*. [Online]. Dordrecht: Springer Netherlands, pp. 3305–3306.
- Royal College of Psychiatrists & Banks** (2007) Challenging behaviour: a unified approach. *British Psychological Society and Royal College of Speech and Language Therapists* **79**, 12–53.
- Ruparelia K, Abubakar A, Badoe E, Bakare M, Visser K, Chugani DC, Chugani HT, Donald KA, Wilmshurst JM, Shih A, Skuse D, and Newton CR** (2016) Autism spectrum disorders in Africa: current challenges in identification, assessment, and treatment: a report on the international child neurology association meeting on ASD in Africa, Ghana, April 3–5, 2014. *Journal of Child Neurology*. [Online] **31**, 1018–1026.
- Sayal K, Prasad V, Daley D, Ford T, and Coghill D** (2018) ADHD in children and young people: prevalence, care pathways, and service provision. *The Lancet Psychiatry*. [Online] **5**, 175–186.
- Stata Corp** (2019) *Stata: Software for Statistics and Data Science. Stata Statistical Software*. [Online]. College Station, Texas: StataCorp LLC. Available from: <https://www.stata.com/> [Accessed: 26 November 2020].
- Strand KB, Chisholm D, Fekadu A, and Johansson KA** (2016) Scaling-up essential neuropsychiatric services in Ethiopia: a cost-effectiveness analysis. *Health Policy and Planning*. [Online] **31**, 504–513.
- Streiner DL, Norman GR, and Cairney J** (2015) *Health Measurement Scales: A Practical Guide to Their Development and use*. [Online]. Oxford University Press. Available from: <https://oxfordmedicine.com/view/10.1093/med/9780199685219.001.0001/med-9780199685219> [Accessed: 17 September 2021].
- Tekola B, Baheretibeb Y, Roth I, Tilahun D, Fekadu A, Hanlon C, and Hoekstra RA** (2016) Challenges and opportunities to improve autism services in low-income countries: lessons from a situational analysis in Ethiopia. *Global Mental Health*. [Online] **3**, e21.
- Tekola B, Kinfe M, Girma F, Hanlon C, and Hoekstra RA** (2020) Perceptions and experiences of stigma among parents of children with developmental disorders in Ethiopia: a qualitative study. *Social Science & Medicine*. [Online] **256**, 113034.
- Tilahun D, Hanlon C, Fekadu A, Tekola B, Baheretibeb Y, and Hoekstra RA** (2016) Stigma, explanatory models and unmet needs of caregivers of children with developmental disorders in a low-income African country: a cross-sectional facility-based survey. *BMC Health Services Research*. [Online] **16**, 152.
- Tomlinson M, Yasamy MT, Emerson E, Officer A, Richler D, and Saxena S** (2014) Setting global research priorities for developmental disabilities, including intellectual disabilities and autism. *Journal of intellectual disability research : JIDR*. [Online] **58**, 1121–1130.
- Wakuma C** (2015) Ethiopian sign language and educational accessibility for the deaf community: a case study on Jimma, Nekemte, Addis Ababa and Hawasa towns. *Journal of Languages and Culture*. [Online] **6**, 9–17.
- Willis GB** (1999) *Cognitive Interviewing: A "How To" Guide*. [Online]. Research Triangle Park, NC: Research Triangle Institute. Available from: <https://www.hkr.se/contentassets/9ed7b1b3997e4bf4baa8d4eceed5cd87/gordonwillis.pdf>.
- World Health Organization** (2001) International classification of functioning, disability and health. [Online]. Available from: <https://apps.who.int/iris/bitstream/handle/10665/42407/9241545429.pdf> [Accessed: 18 February 2020].
- World Health Organization** (2013) *Investing in Mental Health: Evidence in Action*. [Online]. Geneva: World Health Organization. Available from: https://apps.who.int/iris/bitstream/handle/10665/87232/9789241564618_eng.pdf?sequence=1&isAllowed=y.
- World Health Organization** (2018) *Mental Health Atlas 2017*. [Online]. Geneva, Switzerland: World Health Organization. Available from: https://www.who.int/mental_health/evidence/atlas/profiles-2017/en/.
- World Health Organization** (2019) International Statistical Classification of Diseases and Related Health Problems. 11th edition. [Online]. Available from: <https://icd.who.int/dev11/f/en#/http%3a%2f%2fid.who.int%2fid%2fentity%2f334423054> [Accessed: 21 July 2021].
- Zhang Z** (2016) Missing data imputation: focusing on single imputation. *Annals of Translational Medicine*. [Online] **4**, 9.