CrossMark

Screening for Mental Health Among Young South African Children: the Use of the Strengths and Difficulties Questionnaire (SDQ)

Claude Ann Mellins[1] · Qi Xu[2] · Danielle F. Nestadt[1] · Justin Knox[3] · Shuaib Kauchali[4] · Stephen Arpadi[3,5] ·
Jane Kvalsvig[4] · Patrick E. Shrout[2] · Leslie L. Davidson[3,5]

Abstract

Background Little is known about the prevalence of child mental health (MH) problems in sub-Saharan Africa, where poverty, HIV, and family disruption increase risk. One barrier is the lack of MH assessment tools lay staff can validly and reliably administer in settings with few MH professionals.

Methods In a South African (SA) peri-urban cohort, we examined psychometric properties of the Strengths and Difficulties Questionnaire (SDQ), a widely used measure of child emotional and behavioral functioning. Data come from a large population-based study ($N = 1581$) of children 4–6 years old. Lay fieldworkers administered the SDQ in isiZulu to caregivers at baseline and 2 years later. Exploratory factor analysis examined whether the established SDQ five-factor structure and Total Difficulties score would be replicated. The psychometric model was tailored for ordinal items, and target factor rotation was used.

Results Total difficulties, emotional symptoms, and prosocial behavior factors were supported, with partial support for conduct problems. Peer relationships and hyperactivity/inattentive subscale items loaded poorly. Subscale Cronbach's alphas ranged from 0.29 (Peers) to 0.62 (emotional). Internal consistency of total difficulties score was acceptable (0.74); 30% scored in the abnormal range on total difficulties, based on UK norms.

Conclusions SDQ scores in our sample suggest young children in SA are at high risk for MH problems. The SDQ, particularly the total difficulties score, may be a useful screening tool in SA. Yet, some subscales did not work in this language and context; if social skills and hyperactivity/inattention are being considered, modification or additional measures may be needed.

Keywords Strengths and Difficulties Questionnaire (SDQ) · Psychometric · South Africa · Child mental health · Task shifting

✉ Claude Ann Mellins
cam14@cumc.columbia.edu

[1] HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute and Columbia University, 1051 Riverside Drive, Unit 15, New York, NY 10032, USA

[2] Department of Psychology, New York University, New York, NY, USA

[3] Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA

[4] Department of Paediatrics and Child Health, Maternal and Child Health Unit, University of KwaZulu-Natal, Durban, South Africa

[5] Department of Pediatrics, College of Physicians and Surgeons, Columbia University Medical Center, New York, NY, USA

Introduction

Recently, increased attention has been paid to the burden of children's mental health problems globally (Patel et al. 2007). Children and adolescents make up nearly one-third of the world's population (UNICEF 2016), and psychiatric disorders represent a high percentage of their health-related burden, including Disability-Adjusted Life Years (World Health Organization 2013). Recent estimates are that one in five children globally has a psychiatric disorder or significant mental health problem (Belfer 2008; Perou et al. 2013; Polanczyk et al. 2015). Childhood mental health problems are of concern, given their association with poor physical and mental health outcomes throughout childhood and into adulthood, including school dropout, peer relationship difficulties, and difficulty meeting health care needs and transitioning to adulthood (Ford et al. 2007; Kessler et al. 2005).

Springer

Although there are nearly 500 million children aged 0 to 18 years in sub-Saharan Africa, little is known about the prevalence of mental health problems among children in the region, particularly young children, many of whom live in poverty in conditions where family disruption, community violence, and national conflicts are common (UNICEF 2016). One of the few reviews of work in this region found a "psychopathology" prevalence of 14.3% among children 5–16 years old (Cortina et al. 2012), with psychopathology broadly defined in terms of mental health problems, psychiatric disorders, behavioral problems, or psychological distress. South Africa (SA) is a country where child mental health problems may be of particular concern, given high rates of poverty and violence, and a legacy of discrimination due to Apartheid and the HIV epidemic. HIV in particular has taken a significant toll on Black South Africans, with staggering numbers of young children orphaned (UNICEF 2013), increasing risk for mental health problems (Rutter 1971, 1979; Doku 2009).

The paucity of mental health research on African children, including SA children, may be due to multiple factors, including the challenge of meeting basic physical needs including food, shelter, and child survival (Williamson 2005). Also, a significant barrier to understanding child mental health in any context is related to assessment. Children typically present with symptoms in different ways from adults and often lack sufficient language skills to describe their experiences. Evaluation of young children is dependent on parents/caregivers, teachers, and other adults and must be viewed within the context of family and social environments (King et al. 2009).

In the poorest communities in SA, assessment challenges are compounded by a dearth of mental health professionals (Lund et al. 2010). The lack of providers poses a barrier to identifying individuals with mental health problems and providing them with treatment. Recently, calls for use of "lay counselors" in the assessment and treatment of mental health problems have increased (Petersen et al. 2012). However, these efforts have been hindered by a lack of translated and validated standardized clinical instruments that can be easily and effectively administered by non-mental health professionals.

One of the most commonly administered brief mental health screening tools for children that can be used by lay professionals globally is the Strengths and Difficulties Questionnaire (SDQ) (Youth in Mind 2012). The SDQ, completed by caregivers, teachers, or older children themselves, measures behavioral difficulties and pro-social strengths in children ages 3 to 16, with clinical cut-off scores to indicate likely mental health problems. An international review of 48 studies on the psychometric properties of the SDQ caregiver and teacher versions for children ages 4 to 12 found sufficient internal consistency, test-retest reliability, inter-rater agreement, and validity of a five-factor structure, supporting its use across contexts (Stone et al. 2010). However, only a few studies have used the SDQ in Africa (e.g., Kashala et al. 2005; Menon et al. 2007), where ethnic and cultural

differences in childrearing and expectations about childhood in general might impact how caregivers perceive development, emotional and behavioral function, and social relationships. Some of these studies suggest that the SDQ total score has adequate psychometric characteristics, but not all of the factors established by Goodman and colleagues have been equally supported. A few studies using confirmatory and exploratory factor analysis have found the peer problem scale items, and some conduct items, do not always load as expected on published factors (Stone et al. 2010). Moreover, these studies suggest that cut-off scores for significant mental health problems based on British or US-based samples may need to be reconsidered as much larger proportions of the children in settings, such as Zambia and Democratic Republic of Congo, fall outside the normal range (Kashala et al. 2005; Menon et al. 2007). It is not clear if this accurately reflects prevalence of mental health problems or indicates socio-cultural differences.

The SDQ had been translated in SA into Xhosa and Afrikaans, but not isiZulu, the dominant language and culture in KwaZulu-Natal, an area of SA with high poverty and HIV rates, and thus potentially high rates of child mental health problems (Department of Health South Africa 2005). Moreover, no SA studies have evaluated the factor structure of the SDQ and its psychometric characteristics. To our knowledge, few studies in any African country have examined the SDQ in children as young as 4 years, nor its use longitudinally as children develop from preschool to school age. The ability of measures to reflect developmental change is important to ongoing monitoring of mental health. Thus, using longitudinal data from a large epidemiological cohort study of children and their caregivers living in KwaZulu-Natal, in an area affected by poverty and HIV, with few health or social resources, we examined the psychometric performance of the SDQ in isiZulu, including construct validity and internal consistency.

Methods

Participants and Procedures

Asenze is a longitudinal epidemiological study that followed a cohort of pre-school aged children into school age. The study was conducted from 2008 to 2012 in five tribal areas with an estimated population of 67,000 in KwaZulu-Natal, SA. The study area, with semi-rural and peri-urban dwellings, has high levels of unemployment and is situated in a province that has had the highest antenatal HIV prevalence in the country (39.5%) (South Africa National Department of Health 2010).

There were two waves of data collection. First, trained fieldworkers conducted a door-to-door survey, identifying households with children aged 4–6 years; 14,425 households were visited, 2049 with eligible children. If written informed consent was obtained ($n = 1787$; 87.2%), a demographic

interview was conducted with a primary caregiver responsible for the child's daily care, and they were invited to participate in a larger assessment of the child and caregiver approximately 2 weeks later at nearby research offices; 1581 (88%) participated in this assessment (wave 1). They were also invited to participate in a second assessment 2 years later when the child was 6–8 years old; 1409 (89% of wave 1 participants) completed wave 2. The assessment included the SDQ, administered to caregivers by lay fieldworkers who were native speakers of isiZulu and bilingual in English and were trained and supervised by a SA child psychologist and a medical doctor.

Ethical Considerations Study procedures were approved by Institutional Review Boards in both SA and the USA. Informed consent was obtained from caregivers for their own and their child's study participation. Separate consent was obtained for focus group discussions, not originally planned (see below).

Measures

Demographics Data were collected on child age and gender, household composition, household assets, income, and caregiver education during the household survey and updated at each interview.

Strengths and Difficulties The SDQ (Goodman 1997) is a brief, behavioral screening questionnaire that has been translated/adapted into multiple languages and validated worldwide (Youth in Mind 2012). Twenty-five items are rated on a 3-point Likert scale (0 = not true; 1 = somewhat true; 2 = certainly true) to assess emotional and behavioral function, resulting in five subscales (emotional symptoms, conduct, peer relationships [peers], hyperactivity/inattentiveness, and prosocial behavior) and a total difficulties score, based on the first 4 subscales. The SDQ was adapted/translated into isiZulu following standard procedures for translation/back-translation, with discussions to resolve discrepancies (Preciago and Henry 1997) and reviewed with the SDQ author (Goodman, personal communication) to ensure adherence to item meaning. Because of structural differences in the isiZulu language and English, focus groups were conducted with SA native isiZulu speakers, and a SA linguist was consulted before the final translation was adopted (available on the SDQ website). All scores from the subscales were computed at both waves using published scoring algorithms in the SDQ website (http://www.sdqinfo.com/py/sdqinfo/c0.py).

Statistical Analyses of SDQ

Analyses were conducted to determine whether the SDQ five-factor structure would be replicated in this context at both waves. We calculated the internal consistency (Cronbach's alpha) for the five original scales: prosocial, conduct, emotional, hyperactivity/inattentive, peers), as well as for total difficulties. We also used exploratory factor analysis (EFA) with two refinements: a psychometric model tailored for items with three ordinal response options and a target factor rotation, attempting to be as close as possible to a pre-specified target pattern (Browne 2001) of the original factor solution.

Both of these were implemented using Mplus (Version 4.7) (Muthen and Muthen 1998–2010). To fit the factor model, we used a weighted least squares estimator with a mean and variance adjusted chi-square statistic (WLSMV). The adequacy of the fit of the model was evaluated using several fit indices: the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA) (Marsh et al. 1996), weighted root mean square residual (WRSR), and a chi-square test for discrepancy between the model and the data. CFI is considered as an "adequate" fit when its value is greater than .9 and as a "good" fit when its value is greater than .95 (Hu and Bentler 1999). For the RMSEA, a value of .05 is thought to indicate close fit, .08 a fair fit, and .10 a marginal fit (Browne and Cudeck 1993). Yu (2002) denoted WRMR statistics typically less than 1.00 as a "good" fit. All analyses were carried out in each wave separately. Additionally, we examined the degree to which the SDQ scale produces stable and consistent results by computing the correlations between responses at two time points.

Results

Demographics Description

To allow comparison of psychometric results from both waves, we limited our analyses to the children with SDQs completed at both waves ($N = 1394$); 702 (50.4%) were boys, 692 (49.6%) girls (Table 1). The mean age at wave 1 was nearly 5 years. The majority of questionnaires (65.5%) were completed by the child's birth mother and father: 21.1% by grandmothers and 13.4% others. Among caregivers who reported on their education (78%), 6% had no formal education, 4.3% had some primary school, 3.4% completed primary school, 13% had some middle school, 8.5% completed middle school, 40.3% had some high school, 20.9% completed high school, 3.6% had some college. Among caregivers reporting on food security (96%), nearly 6.1% reported food insecurity for over 5 days in the past month and 19.3% reported food insecurity problems for 1–5 days; 74.6% had not experienced food insecurity in the previous month.

Descriptive Statistics of SDQ Items

Table 2 presents descriptive statistics for SDQ items. As indicated, prosocial scale items had higher means and less

Table 1 Demographic characteristics of children in wavec1 (*N* = 1394)

Age in months	Mean	SD
	59.1	7.0
	N	*%*
Gender		
Males	702	50.4
Females	692	49.6
Caregiver type		
Biological mother	887	63.6
Biological father	26	1.9
Grandmother	294	21.1
Others (e.g., relative, sibling)	168	12.0
Missing	19	1.4
Highest education level completed by caregivers		
None	65	4.6
Grade 6 or below	153	11.0
Grade 7 to grade 9	247	17.7
Grade 10 to grade 12	584	41.9
College	39	2.8
Unknown or missing	306	22.0
Days in last month of food insecurity		
None	996	71.5
1–5 days	258	18.5
> 5 days	81	5.8
Missing	59	4.2

variation than those items reflecting difficulties across the two waves.

Internal Consistency

Cronbach's alphas for the scales were consistent at each wave (1, 2): total difficulties (0.74, 0.74), emotional (0.62, 0.62), prosocial (0.57, 0.57), conduct (0.51, 0.47), hyperactivity/inattentive (0.42, 0.54), and peers (0.38, 0.29). Although the values for the total difficulties are acceptable, and values for emotional and prosocial are almost acceptable, the values for conduct, hyperactivity/inattentive, and peers are generally unacceptable. The factor analysis can give more insight into which items are inconsistent.

Factor Analysis

We extracted five factors and specified a target pattern for the items based on Goodman (1997): (1) prosocial, (2) conduct, (3) emotional, (4) hyperactivity/inattentive, and (5) peers. The fit of the five factor model was adequate in both wave 1 ($\chi^2(185) = 557.11$; CFI = .95; TLI = .93; RMSEA = .04; WRMR = 1.01) and wave 2 ($\chi^2(185) = 357.49$; CFI = .98;

Table 2 Mean scores and standard deviation for caregiver SDQ by wave (*N* = 1394)

The 25 SDQ items arranged in the five scales[a]	Wave 1		Wave 2	
	Mean	SD	Mean	SD
Prosocial behavior				
1: Considerate of other people's feelings	1.74	.58	1.87	.45
4: Shares readily with other children	1.73	.62	1.79	.56
9: Helpful if someone is hurt	1.42	.69	1.80	.52
17: Kind to younger children	1.86	.48	1.89	.43
20: Often volunteers to help others	1.75	.58	1.70	.66
Conduct problems				
5: Often has temper tantrums or hot tempers	1.21	.86	1.04	.92
7: Generally obedient (reversed)	.20	.51	.19	.51
12: Often fights with other children	.70	.84	.49	.82
18: Often argumentative with adults	.84	.86	.89	.91
22: Can be spiteful to others	.34	.63	.23	.55
Emotional symptoms				
3: Often complains of headaches	.92	.84	.65	.82
8: Many worries	.35	.66	.36	.65
13: Often unhappy, downhearted	.60	.80	.58	.83
16: Nervous or clingy in new situation	.84	.83	.70	.86
24: Many fears, easily scared	.94	.83	.88	.89
Hyperactivity/inattention				
2: Restless, overactive	1.61	.65	1.59	.70
10: Constantly fidgeting or squirming	1.45	.74	1.46	.79
15: Easily distracted, concentration wanders	.97	.83	1.00	.90
21: Can stop and think things out before acting (reversed)	.86	.76	.59	.74
25: Sees tasks through to the end (reversed)	.83	.75	.50	.69
Peer relationships				
6: Rather solitary, tends to play alone	.79	.92	.64	.88
11: Has at least one good friend (reversed)	.56	.84	.55	.83
14: Generally liked by other children (reversed)	.05	.29	.05	.27
19: Picked on or bullied	1.12	.81	1.03	.88
23: Gets on better with adults than with other children	.51	.73	.42	.71

[a] Each item rated on a 3-point scale (0 to 2)

TLI = .96; RMSEA = .03; WRMR = .80). As shown in Table 3, items generally had loading of .3 or higher on the expected factors. The exceptions to this rule were the peer items, which did not form a coherent factor, and two of the hyperactivity/inattentive items. The two hyperactivity/inattentive items (21 and 25) loaded on the prosocial factor with negative loadings. These were reverse-coded items designed to measure inattention. Another reverse-coded item that loaded with prosocial was item 7 ("generally obedient"), which had a − .5 loading, but it also had a loading of .3 with conduct. In addition to the items that cross-loaded with

Table 3 Five-factor exploratory factor analysis of caregiver SDQ by wave

The 25 SDQ items arranged in the five scales	Wave 1					Wave 2				
	PB	CP	ES	HP	PR	PB	CP	ES	HP	PR
Prosocial behavior										
1: Considerate of other people's feelings	.73	− .03	− .13	.05	− .07	.70	− .14	.03	.01	− .06
4: Shares readily with other children	.50	− .20	.07	− .13	− .12	.50	− .20	− .13	− .06	.01
9: Helpful if someone is hurt	.52	− .03	− .05	.34	.05	.63	.01	− .02	.20	.01
17: Kind to younger children	.67	− .20	.12	− .10	.06	.76	− .04	− .06	.02	− .07
20: Often volunteers to help others	.57	.04	.00	.11	− .05	.55	− .09	.13	− .08	.04
Conduct problems										
5: Often has temper tantrums or hot tempers	.11	.50	.09	− .19	.59	.09	.68	.05	− .06	.63
7: Generally obedient (reversed)	− .52	.30	− .02	.12	− .01	− .52	.31	.02	.10	− .03
12: Often fights with other children	− .10	.53	.03	.15	− .02	− .02	.41	.08	.12	.03
18: Often argumentative with adults	.04	.61	− .02	.05	− .11	.12	.52	− .02	.12	− .24
22: Can be spiteful to others	.03	.61	.03	− .13	− .37	− .02	.65	− .10	− .13	− .25
Emotional symptoms										
3: Often complains of headaches	.03	.08	.38	.03	.06	.03	− .04	.50	.05	.04
8: Many worries	.01	.20	.48	− .07	.09	.01	− .14	.71	− .14	.05
13: Often unhappy, downhearted	− .08	.18	.37	.11	.17	− .03	− .08	.74	.02	.15
16: Nervous or clingy in new situation	− .02	− .10	.80	.03	− .04	.02	.17	.43	.03	− .10
24: Many fears, easily scared	.07	.04	.56	.16	.02	.15	.00	.60	.15	− .17
Hyperactivity/inattention										
2: Restless, overactive	.08	.08	.00	.77	.15	.11	− .02	.00	.82	.14
10: Constantly fidgeting or squirming	.06	− .01	.07	.82	.16	.13	− .05	.02	.90	.18
15: Easily distracted, concentration wanders	− .08	.08	.56	.21	− .12	.05	.24	.17	.33	− .17
21: Can stop and think things out before acting (reversed)	− .48	− .02	.05	− .14	.02	− .34	.04	− .01	.30	− .18
25: Sees tasks through to the end (reversed)	− .38	.00	.07	.04	.05	− .28	.21	.04	.16	− .27
Peer relationships										
6: Rather solitary, tends to play alone	.04	.09	.15	.12	.92	.03	.30	.13	.26	.57
11: Has at least one good friend (reversed)	− .39	− .09	.03	.39	.23	− .34	− .07	.02	.47	.08
14: Generally liked by other children (reversed)	− .63	.04	− .07	− .04	.03	− .66	.03	.27	− .09	.12
19: Picked on or bullied	.18	.16	.20	.23	.02	.14	.13	.38	.05	− .05
23: Gets on better with adults than with other children	.13	.17	− .08	.32	.05	.03	.03	.08	.04	.07

Note: factor loadings over .35 appear in italics

prosocial, item 5 from conduct ("often has temper tantrums") had a high loading with peers in both waves.

Considering the poor performance of hyperactivity/inattentive and peer subscales, we set them aside and fit a 3-factor model to the items in the other three original subscales, prosocial, conduct, and emotional. The fit indices showed that the refined model had a good fit in wave 1 ($\chi^2(63) = 201.06$; CFI = .95; TLI = .92; RMSEA = .04; WRMR = .97) and wave 2 ($\chi^2(63) = 160.76$; CFI = .96; TLI = .94; RMSEA = .03; WRMR = .87) and the majority of items loaded on the expected factors, except for two conduct items. In addition to the conduct factor, item 5 "Often has temper tantrums or hot tempers" and item 7 "Generally obedient" loaded on emotional and prosocial separately. These two items cross-loaded more at wave 2 when the children were older.

The target rotation allowed the latent variables to be correlated rather than being constrained to be independent. In both waves, conduct was positively correlated with emotional ($r = .37, .40$) and negatively correlated with prosocial ($r = − .28, − .35$). These factor-based correlations were somewhat larger than the correlations among the scales formed as simple sums of item responses. Conduct item sums and emotional item sums were positively correlated at .31 and .33, whereas conduct and prosocial sums were negatively correlated at − .18 and − .22 in wave 1 and wave 2, respectively. In both waves, there was almost no correlation between emotional and prosocial regardless of whether the scales were formed as the latent variables (− .05, − .13) or simple item sums (− .05, − .09) (see Table 4).

Table 4 Correlations among prosocial, emotional, and conduct subscales for caregiver SDQ by wave

	Wave 1			Wave 2		
	PB	CP	ES	PB	CP	ES
Prosocial behavior		− .28**	− .05		− .35**	− .13**
Conduct problems	− .18**		.37**	− .22**		.40**
Emotional symptoms	− .05	.31**		− .09**	.33**	

Top-right diagonal values are from EFA; bottom-left diagonal values are based on the average of the 5 items

**p<0.01

The most well-validated and widely used SDQ scale is total difficulties, which summarizes all negative emotional and behavioral items (i.e., excludes prosocial items). Although we know the SDQ has a multifactorial structure, we examined a one-factor solution to determine if all of the items were related to a single difficulties latent dimension. Across both waves, only 13 items, including 3 hyperactivity/inattentive items and 3 peer items, consistently had loadings over .35. The other 7 items did not have high loadings on the factor. When computed based on all 20 items, Cronbach's alpha was 0.74 in both waves; based on 13 items, the alpha was 0.76 at wave 1 and 0.74 at wave 2. The 20-item and 13-item versions of total difficulties had a modest but statistically significant correlation over waves (r (20 items) = .32, r (13 items) = .27). There was no evidence that the more homogeneous set of 13 items had higher stability over time.

Prevalence of Psychological Risk

We compared the distribution of the SDQ scores in KwaZulu-Natal to UK norms given by Goodman (1997) (Tables 5 and 6 and Fig. 1). *Asenze* participants had higher scores on total difficulties and all of its four subscales. For example, 36.7% of our participants had total difficulties scores greater than the UK cutpoint of 17 in wave 1, compared to 9.8% of UK children who received scores above this value (Meltzer et al. 2000). The proportion of children above this value in wave 2 (28.3%) was also elevated relative to UK norms, but less than that of wave 1. Across two waves, 202 children (14.5%) consistently had total difficulties scores above the UK

cutpoint of 17. In general, the mean scores in wave 2 were consistently lower than wave 1 (Tables 5 and 6), while still being larger than UK norms.

Post Hoc Qualitative Interviews and Findings Concerning Caregiver Perceptions of Items

During wave 2, we conducted two focus group discussions (not originally planned) with caregivers to gain socio-cultural insights into the high scores observed on several SDQ subscales, as well as concerns about the peers subscale. We randomly sampled from caregivers of children who had high SDQ scores, from different tribal areas, comprising two groups of six to seven participants. Questions concerned: (1) what defines an 'unhappy child'; (2) what is considered "appropriate" behavior in young children; (3) challenging behaviors; (4) the role of friends for young children; (5) children's recognition of others' feelings and sharing; and (6) obedience and following instructions. A bilingual (isiZulu and English) ethnography fieldworker recruited and consented participants, facilitated the audiotaped discussions, and transcribed and translated the discussions. Two SA team-members extracted the predominant responses to each question.

Caregivers described manifestations of "unhappiness" as refusal to play with other children, eat, and/or talk, and bullying/fighting with others or being bullied. "Crying without a reason" was described as "not normal" and some attributed this to physical pain. Challenging behaviors for caregivers included children being "rude" or disrespectful to caregivers, using vulgar language and fighting with caregivers.

Table 5 Wave 1 scores on each subscale and total difficulties scores based on caregiver-rated SDQ ($n = 1394$)

Five subscales and total difficulties scores	Mean	SD	Median	P25	P75	P90	UK-based cutoffs	Prevalence in wave 1 based on UK cut-offs
Emotional symptoms	3.65	2.51	3	2	5	7	5	23.2%
Conduct problems	3.29	2.19	3	2	5	6	4	27.3%
Hyperactivity/inattention	5.72	2.06	6	4	7	8	7	21.7%
Peer relationships	3.04	2.02	3	4	7	8	4	23.5%
Total difficulties	15.70	6.27	15	11	20	24	17	36.7%
Prosocial behavior	8.50	1.78	9	P75% = 10	P25% = 8	P10% = 6	4	2.2%

Table 6 Wave 2 scores on each subscale and total difficulties scores based on caregiver-rated SDQ ($n = 1394$)

Five subscales and total difficulties scores	Mean	SD	Wave 2 vs. wave 1 (T value)	Median	P25	P75	P90	UK-based cut-offs	Prevalence in wave 2 based on UK cut-offs
Emotional symptoms	3.17	2.56	5.95**	3	1	5	7	5	18.9%
Conduct problems	2.84	2.16	6.90**	2	1	4	6	4	19.9%
Hyperactivity/inattention	5.15	2.28	7.48**	5	4	7	8	7	16.7%
Peer relationships	2.68	1.92	5.09**	2	1	4	6	4	15.9%
Total difficulties	13.84	6.34	9.39**	13	9	18	23	17	28.3%
Prosocial behavior	9.05	1.58	− 10.59**	10	P75% = 10	P25% = 8	P10% = 7	4	1.4%

**p<0.01

The discussion of childhood friendships had relevance for the SDQ Peer subscale. Some caregivers thought children should play mainly with children who resided in their household or family. This allowed for better supervision and also avoided unwanted negative influences "from outside," such as disobedience, bad language, sexual experimentation and smoking, and reduced risk of physical dangers, such as snakebites and sexual abuse. Playing with friends at school under teacher supervision seemed acceptable, but caregivers were more cautious about allowing children to visit school friends outside school. Several caregivers reported that older siblings often looked after younger ones and that caregivers did not usually participate in children's play. Thus, they may not have sufficient ability to report on peer relationships.

There was agreement on teaching children to be helpful and to undertake household chores from a young age. Caregivers emphasized the need to ensure obedience even if it meant occasionally threatening or punishing children. Not being attentive or not completing tasks was attributed to forgetfulness or being distracted by play.

Discussion

In countries where many children grow up in poverty, the ability to identify childhood mental health problems is a critical public health challenge. International best practices for children under the age of 10 are to obtain reports from caregivers. In contexts with limited resources for professional assessments, the likelihood of obtaining useful information from caregivers is increased when using standardized survey instruments that have been used in similar settings and validated against professional assessments. However, standardized measures may not automatically yield precise information when adapted for a new cultural group and translated into a new language. For this reason, we examined carefully the
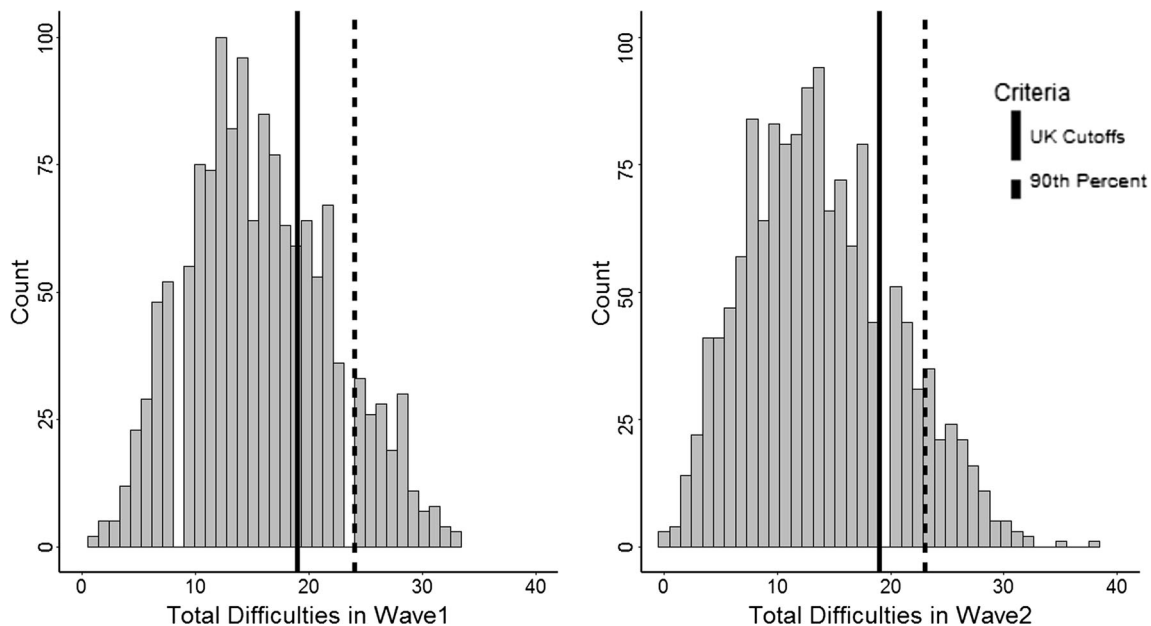

Fig. 1 Total difficulties distribution with 90th percentile cut-off and UK cut-off by wave

psychometric properties of the widely used SDQ after it was translated into isiZulu and used in a multi-wave assessment of a large sample of children in KwaZulu-Natal, SA. To our knowledge, this is one of the few large studies in sub-Saharan Africa to longitudinally follow a young cohort using the SDQ.

In this new context, we confirmed that the emotional sub-scale of the SDQ had adequate internal consistency and evidence of construct validity and that two others (prosocial and conduct) had marginally acceptable internal consistency. Psychometric problems were found in the remaining two sub-scales. The inattention items within the hyperactivity/inattentive subscale did not correlate with the hyperactivity items, and none of the items in the peers subscale showed the expected psychometric structure. Despite the mixed findings for the subscales, we found that the total difficulties score had an acceptable internal consistency of 0.74.

Based on the psychometric analyses, we conclude that the total difficulties score is more useful than the subscales in this language and context. The distribution of the SDQ Total Difficulties scores suggests that young children in our study were at high risk for mental health problems. This risk may be due to multiple factors including poverty, HIV/AIDS, family disruption, and other variables associated with the legacy of Apartheid on Black South Africans. The risk associated with individual children seemed to change over the 2-year follow-up period in our study. The wave 1 and wave 2 total difficulties scores were only correlated 0.23 with each other, with the mean level of problems significantly declining over time. Young children's mental health can change over time, particularly in high stress environments where adverse events can come and go, and thus, we do not necessarily expect high levels of consistency in this context of high poverty and negative familial events.

We were not able to reproduce the 5-factor structure in this context, finding uneven support of the subscales. We identified only one previous study from an African setting that examined the SDQ factor structure, but using teacher-ratings, not caregiver-ratings, and primary school children (7–9 years) (Kashala et al. 2005). The similarities between our studies are interesting, despite the differences in raters. Their factor structure revealed a very similar overall pattern with adequate loadings for the prosocial, emotional and conduct subscales, fair loadings on hyperactivity/inattentive, and poor loadings on peers. In our study, further support for some of the SDQ scales was reflected in the inter-correlations. For example, conduct was negatively associated with prosocial as one would expect. It was also positively associated with emotional, both reflecting behavioral difficulties.

Similar to other studies in Africa, the peers subscale performed the poorest, with items not loading as expected on published factors (Stone et al. 2010). Thus, it may be the least useful in African settings, where, as suggested by our

qualitative data, significant concerns about safety and other cultural factors related to childrearing may result in limitations on child interactions with peers, as well as fewer opportunities for caregivers to actually observe interactions with peers. Also, cultural views on what are appropriate and inappropriate parent-child interactions may vary across LMIC and high-income countries and lead to different ratings on the SDQ items related to conduct and attention, which were derived in a western-European context with different approaches to parenting and caregiving.

The hyperactivity/inattentive subscale did not perform as expected in both waves, seeming to comprise two separate factors. Also, in wave 2 among 6–8-year olds, some of the conduct items loaded on the hyperactivity/inattentive factor. However, the two factors that did emerge do reflect recent literature indicating that attention deficits may or may not co-exist with hyperactivity (American Psychiatric Association 2013). Thus, the two are now often considered separate constructs. Further work in the African context on this particular diagnosis might be helpful for enhancing the SDQ.

Few studies from LMIC have examined use of the SDQ in children as young as 4 years (e.g., Du et al. 2008; Owen et al. 2015), and only Du et al. (2008) had a sample as large as ours, but they did not look at the youngest children separately. Thus, studies with comparison groups for preschool children in LMIC are limited. The 90th percentile cut-offs in our 4–6-year olds were higher on all subscales and the total score compared to the published UK cut-offs, perhaps indicating a greater proportion of children with behavior problems in settings with socioeconomic, health, and historic challenges.

At wave 2, compared to wave 1, the 90th percentile cut-offs of the emotion, conduct, and hyperactivity subscales were the same, and peers and total difficulties were lower. But all sub-scales and total difficulties remained higher for our sample, compared to the UK cut-offs. The higher 90th percentile scores across both waves for total difficulties could reflect more frequent psychosocial problems in this context, or it could reflect cultural differences in understanding of items or perceptions of behavior.

There were several limitations to the study. Adapting an instrument based on the English language to a language with a different structure raises concerns around validity and reliability. Our extensive translation efforts were an attempt to overcome this. However, we did not have the resources for a validation of the SDQ against a clinical assessment by a mental health professional (nor are there many bilingual professionals in SA). Although a large epidemiological study, the data may not reflect other ethnic groups or other countries in Africa. Another limitation is that we only had caregiver reports. The SDQ requires parents to make relatively nuanced judgements about their child. If the child is not causing trouble at home, they may not notice inattention or peer problems.

Thus, multiple raters are often needed to fully understand what is happening with the child, including the child's report as they age, which is possible with the SDQ once children are age 11.

That said, the major strengths of this study were the large population-based sample, the use of longitudinal data with a high retention rate, and a strong analytic approach to the psychometric properties in this context. Our results supported the psychometric structure of the emotional, prosocial, and conduct subscales of the SDQ in SA, suggesting that the overall translation and implementation of the measure in this context was successful. The fact that the psychometric results for hyperactivity/inattentive and peers were not as expected is an important reminder that one cannot assume that standardized measures developed in one cultural context will automatically transfer to other cultural contexts, even if the measures have been widely used in other studies. Just as the basic scientific findings in medicine and epidemiology warrant constant review and critique (Ioannidis 2005), so does the expectation that measures can be easily adapted. The similarity of our findings on peers to other studies, and the difficulties with hyperactivity/inattentive and conduct, suggests that there is a need for modification of the SDQ in this context and development of other tools to assess mental health in young children, particularly if social skills and hyperactivity/inattention are being considered. Given that recent studies highlight the global burden of attention-deficit/hyperactivity disorder (ADHD) in children (Erskine et al. 2014), as well as the importance of social skills and relationships at all stages of development (Dunkel Schetter 2017), assessment that is appropriate to the social and cultural context of both these variables may be critical to fully addressing the mental health challenges of young people.

Fortunately, the data suggest that the SDQ can still be a useful screening tool with young children in this part of SA, with a few caveats. The total difficulties score had good psychometric properties suggesting that it can be used in SA. Standardized mental health tools are important, as they not only provide a useful clinical tool, but allow for cross-cultural comparisons to further understand the need for mental health resources across the globe. Moreover, accurate assessments of prevalence in specific populations can be critical to targeting resources, particularly when there are limited resources for mental health treatment. However, we recommend that tools such as the SDQ not be used as the sole input for clinical decision-making; decisions need to be considered carefully so that children are evaluated correctly given the context they are in and receive appropriate care.

Compliance with Ethical Standards

Ethical Considerations   Study procedures were approved by Institutional Review Boards in both SA and the USA. Informed consent was obtained from caregivers for their own and their child's study participation. Separate consent was obtained for focus group discussions.

Ethical Approval   All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent   Informed consent was obtained from all individual participants included in the study.

Conflict of Interest   The authors declare that they have no conflict of interest.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, D.C: American Psychiatric Association.

Belfer, M. L. (2008). Child and adolescent mental disorders: the magnitude of the problem across the globe. *Journal of Child Psychology and Psychiatry, 49*(3), 226–236.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*(1), 111–150.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Beverly Hills: Sage.

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 & Adolescent Medicine, 166*(3), 276–281.

Department of Health South Africa. (2005). *National HIV and syphilis prevalence survey South Africa 2005*. Pretoria, South Africa.

Doku, P. N. (2009). Parental HIV/AIDS status and death, and children's psychological wellbeing. *International Journal of Mental Health Systems, 3*(1), 1.

Du, Y., Kou, J., & Coghill, D. (2008). The validity, reliability and normative scores of the parent, teacher and self report versions of the strengths and difficulties questionnaire in China. *Journal of Child and Adolescent Psychiatry and Mental Health, 2*(8), 1–15.

Dunkel Schetter, C. (2017). Moving research on health and close relationships forward—A challenge and an obligation: Introduction to the special issue. *American Psychologist, 72*(6), 511.

Erskine, H. E., Ferrari, A. J., Polanczyk, G. V., Moffitt, T. E., Murray, C. J., Vos, T., et al. (2014). The global burden of conduct disorder and

attention-deficit/hyperactivity disorder in 2010. *Journal of Child Psychology and Psychiatry, 55*(4), 328–336.

Ford, T., Collishaw, S., Meltzer, H., & Goodman, R. (2007). A prospective study of childhood psychopathology: Independent predictors of change over three years. *Social Psychiatry and Psychiatric Epidemiology, 42*(12), 953–961.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124.

Kashala, E., Elgen, I., Sommerfelt, K., & Tylleskar, T. (2005). Teacher ratings of mental health among school children in Kinshasa, Democratic Republic of Congo. *European Child & Adolescent Psychiatry, 14*(4), 208–215.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593–602.

King, R., Schwab-Stone, M., Thies, A., Peterson, B., & Fisher, A. (2009). Psychiatric examination of the infant, child & adolescent. In B. Sadock, V. Sadock, & P. Ruiz (Eds.), *Kaplan & Sadock's comprehensive textbook of psychiatry* (9th ed., pp. 3366–3399). Philadelphia: Lippincott Williams & Wilkins.

Lund, C., Kleintjes, S., Kakuma, R., Flisher, A. J., & Consortium, M. H. R. P. (2010). Public sector mental health systems in South Africa: Inter-provincial comparisons and policy implications. *Social Psychiatry and Psychiatric Epidemiology, 45*(3), 393–404. https://doi.org/10.1007/s00127-009-0078-5

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 315–353). Hillsdale: Lawrence Erlbaum Associates, Inc..

Meltzer, H., Gatward, R., Goodman, R., & Ford, F. (2000). *Mental health of children and adolescents in great Britain*. London: The Stationery Office.

Menon, A., Glazebrook, C., Campain, N., & Ngoma, M. (2007). Mental health and disclosure of HIV status in Zambian adolescents with HIV infection: Implications for peer-support programs. *Journal of Acquired Immune Deficiency Syndromes, 46*(3), 349–354.

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus user's guide*. Los Angeles: Muthén & Muthén.

Owen, J. A. P., Mirza, K. A. H., & Goodman, R. (2015). Validating screening tool in Malayalam for mental disorders. *The Indian Journal of Pediatrics, 82*(7), 595–600.

Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: A global public-health challenge. *The Lancet, 369*(9569), 1302–1313.

Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., ... & Parks, S. E. (2013). Mental health surveillance among children—United States, 2005–2011. *MMWR Surveillance Summary, 62*(Suppl 2), 1–35.

Petersen, I., Lund, C., Bhana, A., & Flisher, A. J. (2012). A task shifting approach to primary mental health care for adults in South Africa: Human resource requirements and costs for rural settings. *Health Policy and Planning, 27*(1), 42–51.

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry, 56*(3), 345–365.

Preciago, J., & Henry, M. (1997). Linguistic barriers in health education and services. In J. G. Garcia & M. C. Zea (Eds.), *Psychological interventions and research with Latino populations* (pp. 235–254). Boston: Allyn and Bacon.

Rutter M (1971) Parent-child separation: psychological effects on the children. *Journal of Child Psychology and Psychiatry, 12*(4), 233–260.

Rutter, M. (1979). Separation experiences: A new look at an old topic. *The Journal of Pediatrics, 95*(1):147–154.

South Africa National Department of Health. (2010). *The 2010 national antenatal sentinel HIV & syphilis prevalence survey in South Africa*. Pretoria, South Africa.

Stone, L. L., Otten, R., Engels, R. C., Vermulst, A. A., & Janssens, J. M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4-to 12-year-olds: A review. *Clinical Child and Family Psychology Review, 13*(3), 254–274.

UNICEF. (2013). *Towards an AIDS-free generation: Children and AIDS–sixth stocktaking report*. New York: UNICEF.

UNICEF. (2016). *The state of the world's children 2016 statistical tables*. Retrieved July 1, 2016 from http://data.unicef.org/resources/the-state-of-the-world-s-children-2016-statistical-tables.html.

Williamson, J. (2005). Finding a way forward: Reducing the impacts of HIV/AIDS on vulnerable children and families. In J. Williamson, G. Foster, & C. Levine (Eds.), *A generation at risk: The global impact of HIV/AIDS on orphans and vulnerable children* (pp. 254–277). Cambridge: Cambridge University Press.

World Health Organization. (2013). *Global health estimates summary tables: DALYs by cause, age, and sex*. Retrieved May 24, 2014 from http://www.who.int/entity/healthinfo/global_burden_disease/GHE_DALY_Global_2000_2012.xls?ua=1.

Youth in Mind. (2012). *SDQ: Information for researchers and professionals about the strengths & difficulties questionnaires*. Retrieved May 20, 2014 from http://www.sdqinfo.com/.

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (unpublished doctoral dissertation). University of California, Los Angeles, California.