



Developing an African youth psychosocial assessment: an application of item response theory

Citation

Betancourt, Theresa S., Frances Yang, Paul Bolton, and Sharon-Lise Normand. 2014. "Developing an African Youth Psychosocial Assessment: An Application of Item Response Theory." *Int. J. Methods Psychiatr. Res.* 23 (2) (January 30): 142–160. doi:10.1002/mpr.1420.

Published Version

doi:10.1002/mpr.1420

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:26529623>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Published in final edited form as:

Int J Methods Psychiatr Res. 2014 June ; 23(2): 142–160. doi:10.1002/mpr.1420.

Developing an African youth psychosocial assessment: an application of item response theory

THERESA S. BETANCOURT¹, FRANCES YANG^{2,3}, PAUL BOLTON⁴, and SHARON-LISE NORMAND⁵

¹ Department of Global Health and Population, Harvard School of Public Health, Boston, MA, USA

² Department of Medicine, Harvard Medical School, Boston, MA, USA

³ Institute for Public and Preventive Health, Georgia Regents University, Augusta, GA, USA

⁴ Center for Refugee and Disaster Response, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

⁵ Department of Health Care Policy, Harvard Medical School & Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA

Abstract

This study aimed to refine a dimensional scale for measuring psychosocial adjustment in African youth using item response theory (IRT). A 60-item scale derived from qualitative data was administered to 667 war-affected adolescents (55% female). Exploratory factor analysis (EFA) determined the dimensionality of items based on goodness-of-fit indices. Items with loadings less than 0.4 were dropped. Confirmatory factor analysis (CFA) was used to confirm the scale's dimensionality found under the EFA. Item discrimination and difficulty were estimated using a graded response model for each subscale using weighted least squares means and variances. Predictive validity was examined through correlations between IRT scores (θ) for each subscale and ratings of functional impairment. All models were assessed using goodness-of-fit and comparative fit indices. Fisher's Information curves examined item precision at different underlying ranges of each trait. Original scale items were optimized and reconfigured into an empirically-robust 41-item scale, the African Youth Psychosocial Assessment (AYPA). Refined subscales assess internalizing and externalizing problems, prosocial attitudes/behaviors and somatic complaints without medical cause. The AYPA is a refined dimensional assessment of emotional and behavioral problems in African youth with good psychometric properties. Validation studies in other cultures are recommended.

Keywords

psychometric measurement; children and adolescents; Africa; Uganda; mental health assessment; item response theory (IRT)

Copyright © 2014 John Wiley & Sons, Ltd.

Correspondence Theresa Betancourt, Department of Global Health and Population, Harvard School of Public Health, Francois-Xavier Bagnoud Center for Health and Human Rights, Boston, MA USA. theresa_betancourt@harvard.edu.

Introduction

More than one billion children worldwide live in areas affected by armed conflict. Of these children, a majority live in sub-Saharan Africa (SSA) (UNICEF, 2009). In addition to the risk of physical harm, war-affected children often experience direct and indirect exposure to violence, disrupted family functioning, damaged social structures, limited access to a range of traditional or formal services, all of which serve to increase risks for mental health problems (Guha-Sapir *et al.*, 2005; Newman, 2003; Patel *et al.*, 2007b; UNICEF, 2009). However, in many low-income countries where conflict is widespread, scarce monetary or human resources are devoted to accurate measurement of mental health problems in children and adolescents, resulting in limited data to support claims about the burden of mental health disorders in young people or the outcomes of intervention research (Jacob *et al.*, 2007; Chisholm *et al.*, 2007; Newman, 2003; Patel *et al.*, 2007a, 2007b; Prince *et al.*, 2007; Saraceno *et al.*, 2007; Saxena *et al.*, 2007).

Currently, most assessments used in global mental health research are based on pre-existing standard measures developed in higher resource settings, with limited generalizability to low and middle income countries (LAMICs) (Betancourt *et al.*, 2009b; Patel, 2003, 2007; Van Ommeren, 2003; Yang *et al.*, 2006). In cross-national comparisons of mental disorders measured using standard measures (Alegria *et al.*, 2009; Law and Liu, 2008; Macksoud and Aber, 1996), dramatic differences in prevalence rates across countries (Macksoud and Aber, 1996; Snyder *et al.*, 1997) suggest that standard diagnostic systems need to be examined for cultural or contextual alignment (Avinger and Jones, 2007; Betancourt, in press; Cheng and Williams, 1986; Eisenbruch, 1991; Good, 1997; Guarnaccia *et al.*, 2003; Lewis-Fernandez *et al.*, 2002). In particular, researchers and practitioners working with war-affected youth underscore that emotional and behavioral responses to war are generally appropriately conceptualized within the social and cultural context (Psychosocial Working Group, 2003). The use of the term “psychosocial” indicates that symptoms can be expressed both in terms of internalized and externalized distress, but also linked to interpersonal and social difficulties. Without close consideration paid to local context and cultural fit of assessments to the target population, measurement validity (and, by extension, evaluation data) may be compromised (Bravo *et al.*, 1993; Cha *et al.*, 2007; Jordans *et al.*, 2008).

To address some of these limitations, some researchers routinely use qualitative data to identify core, local conceptualizations of mental health problems (Betancourt *et al.*, 2011a, 2011b; Bolton, 2001; Bolton *et al.*, 2004) in order to select and adapt standard measures or develop new measures for specific populations (Betancourt *et al.*, 2009b; Bolton and Tang, 2002; Child Rights International Network, 2001). For instance, in Zimbabwe, Patel *et al.* (1997) used qualitative methods to identify indigenous symptoms indicative of mental disorders, which were used to create a locally-valid Shona Symptom Questionnaire based on the standard Self Report Questionnaire (Patel *et al.*, 1997; Pynoos and Nader, 1988). In northern Uganda, Betancourt *et al.* (2009a) used qualitative data on common mental health syndromes and their symptoms to develop the Acholi Psychosocial Assessment Instrument (APAI), which was later used in a randomized controlled trial for the treatment of depression (Bolton *et al.*, 2007). A blending of qualitative and quantitative methods was

used to develop the Chinese Health Questionnaire (Cheng and Williams, 1986) and the Primary Care Psychiatric Questionnaire developed in India (Srinivasan and Suresh, 1990).

While there is clearly growing interest in cross-cultural mental health measurement using a mix of qualitative and quantitative methods, few of the measures used for mental health assessment in African children and adolescents have been subjected to rigorous validity testing or psychometric evaluation (Betancourt *et al.*, 2009a; Patel *et al.*, 1997). Building on preliminary studies which developed the APAI in northern Uganda (Betancourt *et al.*, 2009a, 2009b), the present study applied item response theory (IRT)-based analyses to reduce and refine the measure and determine: (a) the dimensionality of its scales; (b) item parameters for each item in each subscale; (c) accuracy of the refined measure for assessing underlying constructs of distress as well as prosocial behavior/attitudes; (d) total information provided by each item on its relevant subscale. As an exploratory step in considering the fit of the refined measure (the African Youth Psychosocial Assessment [AYPA]) to other populations, final items were cross-referenced to qualitative data on expressions of common mental health problems in two other countries in SSA: Rwanda and Sierra Leone.

Materials and methods

The present study was conducted among Acholi Luo youth aged 14–17 living in Unyama and Awer internally displaced person (IDP) camps in Gulu, northern Uganda. During the period of the research (2005), an estimated 1600 children (aged < 18) were living in these two camps. The most commonly spoken language in this region is Acholi Luo. In previous stages of the research (briefly described later), study investigators used qualitative data collected in this setting to develop and validate an APAI in the local Luo language for use with children and adolescents. The present paper presents the IRT-based analysis used to shorten and refine the original APAI scale into a more generalizable AYPA.

Overview of the development of the Acholi Psychosocial Assessment Instrument (APAI)

Qualitative data collection—As discussed in prior publications, 45 free listing interviews and 57 key informant interviews were conducted with youth aged 10–17 and their caregivers to identify common local mental health syndrome terms and associated symptoms, as well as prosocial behaviors and attitudes exhibited by war-affected youth (Betancourt *et al.*, 2009b). Key informants were identified by people in the community as community members who were particularly knowledgeable about each of these mental health problems either because of their connections within the community or because people consulted them about the problem. Here, “community member” excluded paid professionals working in the community but not from the community, such as health workers or other persons with outside training because the intent was to get at locally recognized and accessible language for youth self-report purposes. These knowledgeable local key informants then provided further detail on local perceptions of common mental health symptom terms which were meant to capture the language and conceptualizations of the local community (Betancourt *et al.*, 2009a, 2009b).

Analysis of participant responses was based on Thematic Content Analysis (Smith, 1992) and included tallying all free listing results and using key informant interviews to identify

problem clusters and their associated cover terms as well as those for prosocial/adaptive attitudes and behaviors. Thematic Content Analysis is a descriptive presentation of qualitative data which entails identifying patterns and common themes, in this case, syndrome terms and their indicators, across multiple interviews.

Qualitative findings (Betancourt *et al.*, 2009a) identified five syndromes (the depression and dysthymia-like problems of *two tam*, *par*, and *kumu*; a mixed anxiety and depression-like syndrome *ma lwor*; and conduct problems *kwo maraco*). A group of prosocial and adaptive behaviors were also identified. Indicators of common syndrome symptoms displayed some overlap with Western psychiatric criteria for mood, anxiety, and conduct disorders, but also included important local indicators of distress or dysfunction. Full description of the qualitative methods used to develop the instrument is available in prior publications (Betancourt *et al.*, 2009a; Bolton *et al.*, 2004). A table of the items on the full 60-item APAI, by subscale, is provided in Appendix A, Table A1. The APAI rating format asked about the frequency of each symptom experienced during the previous week. A four-point Likert scale ranging from never (0) to a lot (3) was used. The assessment was designed to be administered verbally due to low literacy among respondents. A visual aid was used to clarify response options.

Validity study—A study to validate the APAI using a comparison of parent and self-reported ratings on presence of local syndrome terms has been published previously (Betancourt *et al.*, 2009a; Bolton, 2001). In brief, purposive sampling via knowledgeable key informants was used to identify Acholi youth (aged 14–17) thought to have one or more of the local syndromes. Target enrollment was 50 cases per syndrome (and 50 cases that demonstrated good levels of prosocial behavior). It was assumed that some young people would manifest multiple syndromes due to comorbidity. “Caseness” for each category was determined through agreement of youth self-report and caregiver report along with duration and severity of local syndrome, because no “gold standard” psychiatric assessment was possible at the time of the study (Betancourt *et al.*, 2009a).

Among the 166 youth in the validity study, the APAI demonstrated satisfactory internal reliability, ranging from $\alpha = 0.70$ – 0.87 for each subscale. Inter-rater reliability and test–retest reliability were good for the internalizing problems subscales, with less robust results on the externalizing problems and prosocial scales. Significant mean differences in scores of self-reported cases versus non-cases were found for all depression-like problems but not for conduct problems nor for anxiety (Betancourt *et al.*, 2009a).

Refinement of the APAI to create the AYPA—For the present analyses we used the full sample of youth screened for the randomized controlled trial (RCT) of depression interventions for Acholi youth (Bolton *et al.*, 2007). In total, $N = 667$ young people (aged 14–17) were referred to the study by local community advisors (community leaders, teachers, community workers, and youth representatives) and screened using the full 60-item APAI measure. Enrollment criteria for screening entailed having lived in the camps for at least one month, and speaking the local language (Acholi Luo). The study design was approved by the Boston Medical Center human subjects committee and the ethical review committee at Uganda's Makerere Institute of Social Research.

In addition to the psychosocial assessment, participants also completed a demographic questionnaire and a measure of functional impairment. The measure of functional impairment was constructed using methods developed by Bolton and Tang (2002) whereby items pertaining to functional impairment were drawn from qualitative data collected in this setting about the activities and responsibilities that children and adolescents must carry out in order to care for/contribute to themselves, their family and community. Questions asked about how much difficulty the respondent had carrying out the day-to-day activities and responsibilities. Response options ranged from having “no more difficulty than peers” (0) to “often cannot do the task” (4).

Data collection—Trained local interviewers administered the full APAI to all youth referred for screening. All assessments were conducted in private following informed consent and youth assent (Bolton *et al.*, 2007). To ensure quality control, study supervisors observed at least 10% of interviews, reviewed all paper surveys, and confirmed data on unavailability or refusal by visiting all adolescents not interviewed.

Statistical methods—Analyses of the RCT screening data among a large sample ($N = 667$) aimed to examine the measure's dimensionality and each item's parameters. Analyses aimed to investigate how well a refined version of the measure would capture latent constructs of distress. To achieve these aims, IRT was selected as the best analytic method for examining underlying latent structures and item properties. IRT models have a number of advantages over classical-test-theory-derived analyses, as they estimate an individual's level of disorder based on both the individual's responses to survey items and on the properties of those items. Traditionally used in educational testing to estimate the latent trait or latent ability of an individual responding to a test item, IRT has also been applied to mental health measurement research to estimate the probability that a given individual characterized by higher or lower levels of latent psychological distress will endorse each possible response option (Chan *et al.*, 2008; Newnham *et al.*, 2012; Yu *et al.*, 2010). IRT parameters include (a) item discrimination – the accuracy with which an item measures the latent trait and (b) item difficulty, which can also be thought of as item severity since the latent trait is distress rather than ability (Deng, 2010; Wang *et al.*, 2000). Examination of these parameters can lead to scale calibrations to ensure that items are scored according to their relation to the latent trait. Additionally, IRT models can also estimate item bias or differential item functioning (DIF) across different subgroups (such as groups distinguished by gender, age, or race) (Banister *et al.*, 2010).

Data analysis—In this study, a multi-step IRT-based structural equation modeling (SEM) approach was used to examine both the latent structure of the scale and the individual observed items that comprise latent constructs (Figure 1). Before beginning, descriptive statistics for the sample were generated in Stata v.11 (StataCorp., 2009), along with the frequency of responses for each item.

In the first step, we used an exploratory factor analysis (EFA) to estimate the dimensionality of the items selected for the APAI scale based on the goodness-of-fit indices. These analyses were conducted in Mplus using the default Geomin rotation (oblique rotation). Second, we dropped all items with loadings less than 0.4, as recommended by procedures for item

banking of emotional distress items by the National Institutes of Health, Patient Reported Outcomes Measurement Information System (Veenema and Schroeder-Bruce, 2002). In a third step, modification indices were examined to identify any further items the removal of which would improve the model fit. The largest model modification indices for each item loading on a different factor were modeled each time until we could no longer improve the fit of the model.

Fourth, the reduced measure was tested for a single underlying latent trait using confirmatory factor analysis (CFA). Mplus was also used for this analysis. It is important to note that in classic IRT analyses, CFA establishes that a single underlying latent trait is present and measurable (Wilamowska *et al.*, 2010). However, in applications of IRT to mental health tools, multiple subdomains of latent distress are frequently present; for example, depression is a unidimensional construct, but it has multiple subfactors, such as dysphoria, interpersonal problems, positive affect, and somatic problems (Ellard *et al.*, 2010). In the case of the APAI, our qualitative data helped us to construct subscales for internalizing (anxiety/depression) and externalizing (aggression/conduct) emotional and behavioral problems along with a scale of prosocial attitudes and behaviors. The use of CFA in the fourth step was important for establishing the level of dimensionality (as opposed to assumed unidimensionality) of the total scale.

Fifth, the IRT item parameters, item discrimination (a) and item difficulty (b), were estimated for the graded response model, similar to a two-parameter logistic (2PL) model for dichotomous data in so far as both discrimination and difficulty (severity) parameters are estimated for each item (Engelhard, 2005). Since there were four response categories (0 = none to 3 = all the time), there were three item difficulty parameters estimated corresponding to each threshold, specified as b_1 , b_2 , and b_3 . These were derived using the weighted least squares means and variances estimator in Mplus software v. 6.0 (Muthén and Muthén, 2010).

Once completed, the model was used to investigate concurrent validity using data on functional impairment. We theorized that underlying emotional distress was likely to correlate with functional impairments (Goldstein and Simpson, 1995), and thus examined correlations between IRT scores (θ) for each mental health problem subscale and ratings of functional impairment. All models were also assessed with goodness-of-fit indices, which included the root mean square error of approximation (RMSEA) (Browne and Cudek, 1993; Muthén *et al.*, 1998) and the comparative fit index (CFI) (Bentler and Chou, 1987; Muthén and Muthén, 2000). The RMSEA provides a measure of discrepancy per model degree of freedom and approaches zero as fit improves. The CFI ranges between zero and one. We rejected models with RMSEA values greater than 0.1 (Browne and Cudek, 1993) and accepted models with values close to 0.06 (models with values of 0.06 or less indicate adequate model fit). Models with CFI values near 0.95 were considered desirable (Bentler, 1990; Muthén, 1989).

Finally, information curves, produced by Mplus, were used to compare the plot of the sum of all the item information functions within each subscale/factor. Information may be thought of as the certainty to which an item measures the underlying construct (Normand *et*

et al., 2006). We examined the information content (as measured by Fisher's Information) for each item on the scale corresponding to each specific unidimensional trait identified at varying levels. This approach allowed us to understand if the precision of different items varied at a higher or lower range of the trait (Normand *et al.*, 2006). Given significant stressors and comorbidity in this war-affected sample, we were interested in identifying items that function well in these populations at a range of underlying distress (that is, at higher levels of θ values).

Upon examining patterns of missingness in the data (less than 10% of items for any subscale), item non-response was assumed to be at random and models were estimated using all available data via full information maximum likelihood (FIML). The final reduced scale was then cross-referenced with qualitative data from Rwanda and Sierra Leone describing depression-like, anxiety-like, and conduct problems among youth. Tallies of items deemed by staff to be “conceptually” equivalent were created following steps employed in measures adaptation research in Rwanda (Betancourt *et al.*, 2012).

Results

Characteristics of the study sample, outlined in Table 1, indicated a mean age of 15 years old (55% female). The mean educational level was 5.1 years of formal schooling. The mean physical functioning score was 15.3 on a scale ranging from 0 to 42. Table 2 shows the item summary with the response frequencies for each question. From 667 respondents, the majority endorsed the “not at all” response for drinking alcohol (96%) and using drugs (98%).

The exploratory factor analyses of the entire item set examined the fit of the three, four or five factor solution. EFA showed that a three-factor model fit the poorest ($\chi^2 = 739$, $df = 272$, $p < 0.001$; CFI = 0.907; RMSEA = 0.051), the four-factor model fit much better with the RMSEA meeting criteria ($\chi^2 = 601$, $df = 273$, $p < 0.001$; CFI = 0.934; RMSEA = 0.042), and finally the five-factor model fit showed a slightly better fit than the four-factor model with the RMSEA meeting criteria, but the CFI still did not meet criteria ($\chi^2 = 534$, $df = 268$, $p < 0.001$; CFI = 0.947; RMSEA = 0.039) and the item groupings were less clinically meaningful. For the four factor solution, the RMSEA was near desirable criteria ($\chi^2 = 601$, $df = 273$, $p < 0.001$; CFI = 0.934; RMSEA = 0.042). After elimination of 19 APAI items (APAI06, APAI09, APAI12, APAI14, APAI22, APAI25, APAI27, APAI30, APAI34, APAI40, APAI45, APAI46, APAI47, APAI48, APAI50, APAI52, APAI56, APAI57, and APAI59) with loadings less than 0.4, the fit for the four-factor CFA model improved ($\chi^2 = 1711.786$, $df = 773$, $p < 0.001$; CFI = 0.942; RMSEA = 0.043) and the factor indicators were meaningfully loaded onto factors that made clinical and theoretical sense. IRT analyses confirmed that a four factor solution – internalizing problems, externalizing problems, prosocial attitudes/ behaviors, and somatic complaints without medical cause – best fit the data and validated the broad domains manifest in our qualitative data. The somatic subscale was unanticipated but logical given qualitative findings on the tendency of individuals to somaticize distress in this setting. The final IRT parameter estimates are presented in Table 3.

In analyses of concurrent validity using correlations with functional impairment, the refined subscale for internalizing emotional and behavioral problems in this setting (Factor 1) was correlated with impairment in functioning at 0.36 ($p < 0.001$). The subscales for externalizing problems (Factor 2) and somatic complaints without medical cause (Factor 4) were also correlated with impairment in functioning at relatively the same level as internalizing problems (0.28, $p < 0.001$ and 0.31, $p < 0.001$, respectively). The subscale for prosocial attitudes and behavior (Factor 3) was not significantly correlated with impairment in functioning (0.07, $p = 0.15$).

Figure 2 shows the test information curves for Factors 1 to 4. Overall, the items in the internalizing factor contributed high discrimination and difficulty parameters to yield high information content and wide distribution across different levels of severity, compared to the other sub-scales. The externalizing/conduct items yielded the next highest discrimination parameters summed together, but covered a narrower and more skewed range based on the difficulty parameter. Although the prosocial items covered a large range of the prosocial trait, total information was lower than the previous two traits mentioned. Finally, the somatic items also covered a wide range of the trait, but yielded the lowest information overall when compared to the other traits.

Figure 3 shows the item characteristic curves for each item within each factor. The items are shown for the probability of endorsing the 0 = none and the 3 = all the time categories across the latent trait levels for each of the following four factors: internalizing problems, externalizing problems, prosocial attitudes and behaviors, and somatic complaints without medical cause. In terms of “difficulty” or in this case, likelihood of endorsement depending on the level of the underlying trait, item characteristic curves for the internalizing and externalizing problems show those on the lower level of the respective traits are close to 100% probability of endorsing the “never” response category, except for item 35 “I can do nothing to help myself,” which was still at 80% probability of endorsing the “never” category at the lowest end of the internalizing problems latent trait level. Interestingly, item 44 “I use drugs” was less endorsed relative to other externalizing problems items among those with higher levels of externalizing problems latent trait.

The internalizing subscale with items such as “I feel sad” shares items in common with other standard measures assessing “internalizing emotional and behavioral problems”, such as the Achenbach Youth Self Report (Achenbach, 1992) and the Strengths and Difficulties Questionnaire (Goodman, 1997), but also contained items more similar to African expressions of distress, such as “I feel a lot of pain in my heart” and “I sit with my cheek in my palm.” The internalizing subscale had the most total information (area under the subscale information curves), followed by the externalizing subscale, the prosocial subscale and the subscale for somatic complaints without medical cause. Estimates for total information are now also included in Tables 4 and 5. In terms of item performance, we observed that for the internalizing problems subscale, the lowest discriminating item was “I feel I can do nothing to help myself” ($a = 0.40$), which means that responses to this question do less well in distinguishing between those with low and high levels of underlying internalizing emotional and behavioral problems. This item was retained because in a displacement camp, one's options for helping him/herself are limited by circumstances, and

because modification indices suggest that dropping it did not improve the scale. However, this item's performance will need to be tested in other settings beyond war-affected ones in order to determine if the item performance is improved in situations where children and youth can exert greater self-efficacy. For internalizing problems, the highest discriminating item was "I have lots of worries" ($a = 0.78$), which means that this item distinguishes well between youth with low versus high levels of underlying internalizing problems. For the externalizing problems subscale, the highest discriminating item was "I insult friends" ($a = 0.90$), whereas items such as "I deceive" were less discriminating ($a = 0.60$). For prosocial attitudes and behavior, the most discriminating item was "I cooperate with others" ($a = 0.73$), while items like "I play together with others" showed poorer discrimination ($a = 0.40$). For somatic complaints without medical cause, the item "I have pain all over my body" showed the best discrimination ($a = 0.76$), while the item "I get headaches" was less discriminating ($a = 0.60$). Seventy-two percent of items on the internalizing and externalizing subscales were also matched to total problem items in qualitative data collected in Rwanda (Betancourt *et al.*, 2011b); 68% of items on the internalizing subscale were matched to internalizing problems identified in qualitative data from Sierra Leone (Betancourt, 2010).

Traditional reliability analyses (Cronbach's alpha) were also performed for each of the refined subscales. While the test is limited as an artifact of the number of items in the scale, all subscales had satisfactory to excellent alpha values: prosocial/adaptive ($a = 0.72$), somatic complaints without medical cause ($a = 0.74$), externalizing problems ($a = 0.83$), and internalizing problems ($a = 0.88$). A copy of the final AYPa measure is provided in Appendix B, Table 1.

Discussion

The present study demonstrates how IRT analyses may be applied to refining and optimizing the psychometric functioning of a qualitatively-derived psychosocial assessment, in this case for use with youth in SSA. This study contributes to a growing body of research on cross-cultural assessment by focusing on youth in SSA – a region where exposure to difficulties such as war and disease are common, but child and adolescent mental health research is limited in part by the lack of psychometrically-sound assessments. Grounded in qualitative data elicited from Uganda, the final refined and validated AYPa measure demonstrates potential for broader applicability to other African settings. Qualitative data on expressions of child mental health problems and prosocial behaviors in Rwanda (Betancourt *et al.*, 2011b) and Sierra Leone (Betancourt, 2010) share many similarities with items in the AYPa's internalizing and externalizing problem scales. In particular, the phrasing of items in the AYPa is more comparable to how mental health problems of children and adolescents are expressed in SSA (Betancourt *et al.*, 2009b 2011b; Dorit, 2001). For instance, items capture expressions of sadness (i.e. having "pain in the heart" and "sitting with cheek in palm"), social withdrawal ("staying away from others") and anxiety or rumination ("thinking too much"), conduct ("being a rough person") and prosocial attitudes/behaviors ("welcoming others") documented in qualitative data from several African settings.

Similarities across several SSA cultural groups indicate that the AYPAs may have broader applicability as a dimensional measure of internalizing and externalizing problems, somatic complaints without medical cause, and prosocial behavior beyond the Ugandan sample examined here. However, study analyses also show that it is important to consider AYPAs item functioning in context, as some items may be less discriminating given ongoing situations of insecurity or high underlying levels of distress or physical illness. Future applications of the measure might combine the subscales on internalizing, externalizing and somatic problems without medical cause to create a “total problems score” for evaluation purposes as well as examining outcomes on particular subscales depending on study goals. Like standard measures used frequently in the United States and the United Kingdom including the ASEBA Youth Self Report and Child Behavior Checklist and the Strengths and Difficulties Questionnaire (Achenbach, 1991a, 1991b; Achenbach *et al.*, 2008; Achenbach and Rescorla, 2001; Goodman, 1997), the prosocial attitudes and behaviors scale may be used to assess more positive developmental outcomes while the problem subscales and total problem score subscales capture elements of psychopathology and distress. Assessments of symptoms are likely to be most informative when paired with measures of functional impairment such as the WHO DAS scales or the locally-derived functioning scale used here (World Health Organization, 2010).

Study limitations must be noted. The finding that the subscale for somatic complaints without medical cause was weakest must be considered in the context of a setting with high disease burden and limited services. In such a situation, even with adequate instructions, participants may find it challenging to distinguish somatic complaints not due to a medical cause from those due to illnesses such as malaria. No independent medical evaluation was able to be conducted to make this determination, thus for this study we are limited to the results of self-report. However, this finding is consistent with other research indicating that somatic symptoms contribute less reliable information than hallmark symptoms of depressed mood and anhedonia in the diagnosis of depression in many non-Western settings (Coulter, 2013; United Nations Millennium Project, 2005). A methodological limitation of Mplus is the lack of item fit statistics that other programs may offer. However, the strong theoretical and methodological decisions prior to IRT analysis has shown through the model fit statistics that the items chosen for the current study are adequate for measuring each of the domains in this scale. In addition, analyses examining traditional test reliability are limited as an artifact of the number of items in the scale which likely has implications for the somatic complaints subscale which is brief at four items following IRT. Overall, the study's reliance on youth self-reports remains an important limitation. In future research on the AYPAs among other African populations, it would be valuable to assess validity using comparison to the ratings of a child and adolescent psychiatrist familiar with the culture and setting. In the end, the AYPAs is not intended to serve as a diagnostic tool for assessing specific mental disorders. Rather, it is intended to assess a broad range of externalizing and internalizing problems relevant to mental health syndrome expressions in SSA as well as prosocial behaviors and attitudes important to healthy social functioning in children and adolescents.

In conclusion, the refined and shortened AYPAs measure is a promising tool for assessing internalizing emotional and behavioral problems and externalizing problems, somatic

complaints without medical cause, and prosocial attitudes and behaviors among youth in SSA. The core items in the AYPAs are not unique to northern Uganda and have potential applications to children, adolescents, and youth in other regions. The measure is currently being piloted in other parts of Uganda as well as among war-affected youth in the Democratic Republic of Congo. Initial results are promising, but more research is needed to examine the performance of the scale across a variety of African settings.

Overall, this study demonstrates that IRT-driven analyses can be applied to the refinement of instruments on emotional and behavioral problems in children derived from qualitative data. The approach taken to develop this instrument and examine its psychometric properties can be replicated in future research to expand the measures available to capture psychosocial adjustment among diverse cultural groups.

Acknowledgments

This analysis was funded by Grant #1K01MH077246-01A2 from the National Institute of Mental Health and Grant P60MD002261 from the National Center for Minority Health and Disparities, and by the François-Xavier Bagnoud Center for Health and Human Rights. Primary data collection was supported by World Vision Uganda and War Child Holland. The authors would like to thank Ista Zahn and Robert Brennan for their help with data analyses. The authors are grateful to the children, youth and families of the Awer and Unyama IDP camps who participated in this research.

Appendix A

Table A1

Symptoms of locally-described syndromes included in the 60-item APAI

Symptom	Two tam	Kumu	Par	Ma Iwor	Kwo Maraco	Pro-social
APAI 1: I listen to others and elders						*
APAI 2: I play together with others						*
APAI 3: I have a lot of thoughts	*		*			
APAI 4: I have constant worries	*					
APAI 5: I have pain all over my body	*					
APAI 6: My brain is not functioning well	*					
APAI 7: I think I am of no use	*					
APAI 8: I think about suicide	*		*			
APAI 9: I talk constantly about my problems	*					
APAI 10: I sit alone	*	*				
APAI 11: I share with others						*
APAI 12: I lose interest in school	*				*	
APAI 13: I get headaches	*	*				
APAI 14: I lose my appetite		*	*	*		
APAI 15: I feel a lot of pain in my heart		*				
APAI 16: I sit with my cheek in my palm		*				
APAI 17: I cry when I am alone		*				
APAI 18: I do not sleep at night		*		*		
APAI 19: I am disobedient		*	*		*	
APAI 20: I feel cold		*				

Symptom	Two tam	Kumu	Par	Ma Iwor	Kwo Maraco	Pro-social
APAI 21: I share food and eat with others						*
APAI 22: I lie down all the time (during the day)		*				
APAI 23: I have lots of worries		*	*			
APAI 24: I want to be alone			*	*		
APAI 25: I am easily annoyed			*			
APAI 26: I hold my head			*			
APAI 27: I lose concentration in class			*			
APAI 28: I drink alcohol			*	*	*	
APAI 29: I insult friends			*			
APAI 30: I don't greet people			*	*		
APAI 31: I help others						*
APAI 32: I don't think straight			*			
APAI 33: I mutter to myself			*			
APAI 34: I don't trust			*			
APAI 35: I feel I can do nothing to help myself			*			
APAI 36: I fight					*	
APAI 37: I use bad language					*	
APAI 38: I am disrespectful					*	
APAI 39: I misbehave					*	
APAI 40: I am disinterested					*	
APAI 41: I welcome others						*
APAI 42: I deceive					*	
APAI 43: I am a rough person					*	
APAI 44: I use drugs (like jayi, marijuana)					*	
APAI 45: I cling to elders				*		
APAI 46: I think I do not have a future				*		
APAI 47: I am constantly running around				*		
APAI 48: I don't like noise				*		
APAI 49: I think people are chasing me				*		
APAI 50: I have a fast heart rate				*		
APAI 51: I cooperate with others						*
APAI 52: I fear being alone				*		
APAI 53: I feel sad	*					
APAI 54: I think of bad things	*					
APAI 55: I am weak	*	*	*			
APAI 56: I don't feel like talking	*	*				
APAI 57: I am forgetful	*					
APAI 58: I cry continuously	*		*			
APAI 59: I do not care whether I live or die	*					
APAI 60: I respect others						*

Note: The presence of an asterisk (*) indicates the inclusion of the named symptom (column) in the APAI item denoted (row).

Appendix B

Table B1

African Youth Psychosocial Assessment (AYPA)

Item	Symptom	Pe	Manok	Madwong	Tutwal
		<i>Never</i>	<i>Somewhat</i>	<i>Often</i>	<i>All the time</i>
AYPA 1	Awinyo lok pa dano mukene ki pa ludito <i>I listen to others and elders</i>	0	1	2	3
AYPA 2	Atuku kacel ki dano mukene <i>I play together with others</i>	0	1	2	3
AYPA 3	Atamo madwong <i>I have a lot of thoughts</i>	0	1	2	3
AYPA 4	Atiko par <i>I worry constantly</i>	0	1	2	3
AYPA 5	Koma weng rem arema <i>I have pain all over my body</i>	0	1	2	3
AYPA 6	Atamo ni konya pe <i>I think I am of no use</i>	0	1	2	3
AYPA 7	Atamo ni a dene <i>I think about suicide</i>	0	1	2	3
AYPA 8	Abedo kena-kena <i>I sit alone</i>	0	1	2	3
AYPA 9	Aribu ki dano mukene <i>I share with others</i>	0	1	2	3
AYPA 10	Wiya bara <i>I get headaches</i>	0	1	2	3
AYPA 11	Awinyo ma lit I cwinya <i>I feel a lot of pain in my heart</i>	0	1	2	3
AYPA 12	Abedo ma nongo ateno tika <i>I sit with my cheek in my palm</i>	0	1	2	3
AYPA 13	Akok ka atye kena <i>I cry when I'm alone</i>	0	1	2	3
AYPA 14	Pe anino I dye wor <i>I do not sleep at night</i>	0	1	2	3
AYPA 15	Ading <i>I am disobedient</i>	0	1	2	3
AYPA 16	Awinyo koyo <i>I feel cold</i>	0	1	2	3
AYPA 17	Aribu cam dok acamo kacel ki dano mukene <i>I share food and eat with others</i>	0	1	2	3
AYPA 18	Aparo par madwong <i>I have lots of worries</i>	0	1	2	3

Item	Symptom	Pe	Manok	Madwong	Tutwal
		<i>Never</i>	<i>Somewhat</i>	<i>Often</i>	<i>All the time</i>
AYPA 19	Amito bedo kena <i>I want to be alone</i>	0	1	2	3
AYPA 20	Amako wiya <i>I hold my head</i>	0	1	2	3
AYPA 21	Amato kongo <i>I drink alcohol</i>	0	1	2	3
AYPA 22	Ayeto luwota <i>I insult friends</i>	0	1	2	3
AYPA 23	Akonyo dano mukene <i>I help others</i>	0	1	2	3
AYPA 24	Pe atamo atir <i>I don't think straight</i>	0	1	2	3
AYPA 25	Ajwat kekena <i>I mutter to my self</i>	0	1	2	3
AYPA 26	Awinyo ni ape kigin mo wek akonye kede <i>I feel I can do nothing to help myself</i>	0	1	2	3
AYPA 27	Alwenyo <i>I fight</i>	0	1	2	3
AYPA 28	Atiyo ki leb marac <i>I use bad language</i>	0	1	2	3
AYPA 29	Pe awor <i>I am disrespectful</i>	0	1	2	3
AYPA 30	Atime atata <i>I misbehave</i>	0	1	2	3
AYPA 31	Ajolo dano mukene <i>I welcome others</i>	0	1	2	3
AYPA 32	Abwolo bwola <i>I deceive</i>	0	1	2	3
AYPA 33	An labwamii <i>I am a rough person</i>	0	1	2	3
AYPA 34	Atiyo ki yat calo (jayi, mairungi) <i>I use drugs (like jayi, marijuana)</i>	0	1	2	3
AYPA 35	Atamo ni dano tye ka ryema <i>I think people are chasing me</i>	0	1	2	3
AYPA 36	Aribe ki dano mukene <i>I cooperate with others</i>	0	1	2	3
AYPA 37	Cwinya cwer <i>I feel sad</i>	0	1	2	3
AYPA 38	Atamo jami maraco <i>I think of bad things</i>	0	1	2	3
AYPA 39	Atiko kok akoka				

Item	Symptom	Pe	Manok	Madwong	Tutwal
		<i>Never</i>	<i>Somewhat</i>	<i>Often</i>	<i>All the time</i>
AYPA 40	<i>I cry continuously</i>	0	1	2	3
	Aworo dano mukene				
AYPA 41	<i>I respect others</i>	0	1	2	3
	Agoro				
	<i>I am weak</i>	0	1	2	3

References

- Achenbach, TM. Manual for the Child Behavior Checklist/4-18 and 1991 Profile. University of Vermont, Department of Psychiatry; Burlington, VT: 1991a.
- Achenbach, TM. Manual for the Youth Self-Report and 1991 profile. University of Vermont, Department of Psychiatry; Burlington, VT: 1991b.
- Achenbach, TM. Taxonomy of internalizing disorders of childhood and adolescence.. In: Reynolds, WM., editor. *Internalizing Disorders in Children and Adolescents*. Wiley; New York: 1992. p. 19-60.
- Achenbach, TM.; Rescorla, LA. Manual for the ASEBA School-age Forms & Profiles. University of Vermont, Research Center for Children, Youth, & Families; Burlington, VT: 2001.
- Achenbach TM, Becker A, Dopfner M, Heiervang E, Roessner V, Steinhausen HC, Rothenberger A. Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*. 2008; 49(3):251–275. DOI: JCPP1867 [pii] 10.1111/j.1469-7610.2007.01867.x [doi]. [PubMed: 18333930]
- Alegria M, Shrout PE, Torres M, Lewis-Fernández R, Abelson JM, Powell M, Interian A, Lin J, Laderman M, Canino G. Lessons learned from the clinical reappraisal study of the Composite International Diagnostic Interview with Latinos. *International Journal of Methods in Psychiatric Research*. 2009; 18(2):84–95. DOI: 10.1002/mpr.280. [PubMed: 19507168]
- Avinger KA, Jones RA. Group treatment of sexually abused adolescent girls: a review of outcome studies. *American Journal of Family Therapy*. 2007; 35(4):315–326. DOI: 10.1080/01926180600969702.
- Banister, J.; Bloom, DE.; Rosenberg, L. *Population Aging and Economic Growth in China, Program on the Global Demography of Aging Working Paper Series*. Harvard School of Public Health; Cambridge: 2010.
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin*. 1990; 107(2):238–246. DOI: 10.1037/0033-2909.107.2.238. [PubMed: 2320703]
- Bentler PM, Chou C-P. Practical issues in structural modeling. *Sociological Methods & Research*. 1987; 16(1):78–117. DOI: 10.1177/0049124187016001004.
- Betancourt TS. A longitudinal study of psychosocial adjustment and community reintegration among former child soldiers in Sierra Leone. *International Psychiatry*. 2010; 7(3):60–62.
- Betancourt, TS. Using mixed methods to plan and evaluate mental health programs for war-affected children in Sub-Saharan Africa.. In: Kleinman, A.; Akyeampong, E., editors. *The Culture of Mental Illness and Psychiatric Practice in Africa*. Indiana University Press; IN, USA: in press
- Betancourt TS, Bass J, Borisova I, Neugebauer R, Speelman L, Onyango G, Bolton P. Assessing local instrument reliability and validity: a field-based example from northern Uganda. *Social Psychiatry and Psychiatric Epidemiology*. 2009a; 44(8):685–692. [PubMed: 19165403]
- Betancourt TS, Speelman L, Onyango G, Bolton P. A qualitative study of mental health problems among children displaced by war in northern Uganda. *Journal of Transcultural Psychiatry*. 2009b; 46(2):238–256. DOI: 10.1177/1363461509105815.
- Betancourt TS, Meyers-Ohki S, Stevenson A, Ingabire C, Kanyanganzi F, Munyana MC, Mushashi C, Teta S, Fayida I, Cyamatara FR, Stulac SN, Beardslee W. Using mixed-methods research to adapt

and evaluate a family strengthening intervention in Rwanda. *African Journal of Traumatic Stress*. 2011a; 2(1):32–45.

- Betancourt TS, Rubin-Smith J, Beardslee WR, Stulac SN, Fayida I, Safren SA. Understanding locally, culturally, and contextually relevant mental health problems among Rwandan children and adolescents affected by HIV/AIDS. *AIDS Care, iFirst*. 2011b:1–12. DOI: 10.1080/09540121.2010.516333.
- Betancourt TS, Scorza P, Meyers-Ohki SE, Mushashi C, Kayiteshonga Y, Binagwaho A, Beardslee WR. Validating the center for epidemiological studies depression scale for children in Rwanda. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2012; 51(12):1284–1292. [PubMed: 23200285]
- Bolton P. Cross-cultural validity and reliability testing of a standard psychiatric assessment instrument without a gold standard. *Journal of Nervous and Mental Disease*. 2001; 189(4):238–242. DOI: 10.1097/00005053-200104000-00005. [PubMed: 11339319]
- Bolton P, Tang AM. An alternative approach to cross-cultural function assessment. *Social Psychiatry and Psychiatric Epidemiology*. 2002; 37(11):537–543. DOI: 10.1007/s00127-002-0580-5. [PubMed: 12395144]
- Bolton P, Wilk CM, Ndogoni L. Assessment of depression prevalence in rural Uganda using symptom and function criteria. *Social Psychiatry and Psychiatric Epidemiology*. 2004; 39(6):442–447. DOI: 10.1007/s00127-004-0763-3. [PubMed: 15205728]
- Bolton P, Bass J, Betancourt TS, Speelman L, Onyango G, Clougherty KF, Neugebauer R, Murray L, Verdeli H. Interventions for depression symptoms among adolescent survivors of war and displacement in northern Uganda: a randomized controlled trial. *Journal of the American Medical Association*. 2007; 298(5):519–527. DOI: 10.1001/jama.298.5.519. [PubMed: 17666672]
- Bravo M, Woodbury-Farina M, Canino GJ, Rubio-Stipec M. The Spanish translation and cultural adaptation of the Diagnostic Interview Schedule for Children (DISC) in Puerto Rico. *Culture, Medicine and Psychiatry*. 1993; 17(3):329–344. DOI: 10.1007/BF01380008.
- Browne, M.; Cudek, R. Alternative ways of assessing model fit. In: Bollen, K.; Long, J., editors. *Testing Structural Equation Models*. Sage; Thousand Oaks, CA: 1993. p. 136-162.
- Cha E-S, Kim KH, Erlen JA. Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing*. 2007; 58(4):386–395. DOI: 10.1111/j.1365-2648.2007.04242.x. [PubMed: 17442038]
- Chan EYY, Griffiths S, Gao Y, Chan CW, Fok TF. Addressing disparities in children's health in China. *Archives of Disease in Childhood*. 2008; 93:346–352. DOI: 10.1136/adc.2007.130146. [PubMed: 18356388]
- Cheng T-A, Williams P. The design and development of a screening questionnaire (CHQ) for use in community studies of mental disorders in Taiwan. *Psychological Medicine*. 1986; 16(02):415–422. DOI: 10.1017/S0033291700009247. [PubMed: 3726013]
- Child Rights International Network. Rwanda Country Report. Better Care Network; New York: 2001.
- Chisholm D, Flisher AJ, Lund C, Patel V, Saxena S, Thornicroft G, Tomlinson M. Scale up services for mental disorders: a call for action. *Lancet*. 2007; 370(9594):1241–1252. [PubMed: 17804059]
- Coulter S. Systemic psychotherapy as an intervention for post-traumatic stress responses: An introduction, theoretical rationale and overview of developments in an emerging field of interest. *Journal of Family Therapy*. 2013; 32(4):381–406.
- Deng, Q. *Interconnection Among Global Problems in Northeast Asia*. Nautilus Institute for Security and Sustainability; San Francisco, CA: 2010. Natural disasters, migration and urban insecurity in China..
- Dorit R-S. Reducing risk for children in changing cultural contexts: recommendations for intervention and training. *Child Abuse & Neglect*. 2001; 25(2):231–248. DOI: 10.1016/s0145-2134(00)00242-8. [PubMed: 11330922]
- Eisenbruch M. From post-traumatic stress disorder to cultural bereavement: diagnosis of southeast Asian refugees. *Social Science & Medicine*. 1991; 33(6):673–680. [PubMed: 1957187]
- Ellard KK, Fairholme CP, Boisseau CL, Farchione TJ, Barlow DH. Unified protocol for the transdiagnostic treatment of emotional disorders: protocol development and initial outcome data. *Cognitive and Behavioral Practice*. 2010; 17(1):88–101. DOI: 10.1016/j.cbpra.2009.06.002.

- Engelhard, G. Encyclopedia of Statistics in Behavioral Science. John Wiley & Sons, Ltd.; UK: 2005. Item response theory (IRT) models for rating scale data..
- Goldstein, JM.; Simpson, JC. Validity: definitions and applications to psychiatric research.. In: Tsuang, MT.; Tohen, M.; Zahner, GEP., editors. Textbook in Psychiatric Epidemiology. Wiley-Liss Inc.; New York: 1995. p. 229-242.
- Good BJ. Studying mental illness in context: local, global, or universal? *Ethos*. 1997; 25:230–248.
- Goodman R. The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry*. 1997; 38(5):581–586. DOI: 10.1111/j.1469-7610.1997.tb01545.x. [PubMed: 9255702]
- Guarnaccia PJ, Lewis-Fernandez R, Marano MR. Toward a Puerto Rican popular nosology: nervios and ataque de nervios. *Culture, Medicine and Psychiatry*. 2003; 27(3):339–366. DOI: 10.1023/A:1025303315932.
- Guha-Sapir D, van Panhuis WG, Degomme O, Teran V. Civil conflicts in four African countries: a five-year review of trends in nutrition and mortality. *Epidemiologic Reviews*. 2005; 27(1):67–77. DOI: 10.1093/epirev/mxi010. [PubMed: 15958428]
- Jacob KS, Sharan P, Mirza I, Garrido-Cumbrera IM, Seedat S, Mari JJ, Saxena S. Mental health systems in countries: where are we now? *Lancet*. 2007; 370:1061–1077. DOI: 10.1016/S0140-6736(07)61241-0. [PubMed: 17804052]
- Jordans MJ, Komproe IH, Ventevogel P, Tol WA, de Jong JT. Development and validation of the child psychosocial distress screener in Burundi. *American Journal of Orthopsychiatry*. 2008; 78(3):290–299. DOI: 10.1037/a0014216. [PubMed: 19123748]
- Law S, Liu P. Suicide in China: unique demographic patterns and relationship to depressive disorder. *Current Psychiatry Reports*. 2008; 10:80–86. [PubMed: 18269899]
- Lewis-Fernandez R, Guarnaccia PJ, Martinez IE, Salman E, Schmidt A, Liebowitz M. Comparative phenomenology of ataques de nervios, panic attacks, and panic disorder. *Culture, Medicine and Psychiatry*. 2002; 26(2):199–223. DOI: 10.1023/A:1016349624867.
- Macksoud MS, Aber JL. The war experiences and psychosocial development of children in Lebanon. *Child Development*. 1996; 67(1):70–88. DOI: 10.2307/1131687. [PubMed: 8605835]
- Muthén B. Latent variable modeling in heterogeneous populations. *Psychometrika*. 1989; 54(4):557–585. DOI: 10.1007/bf02296397.
- Muthén BO, Muthén LK. The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. *Journal of Studies on Alcohol and Drugs*. 2000; 61(2):290–300.
- Muthén, LK.; Muthén, BO. *Mplus User's Guide*. 6th edition. Muthén & Muthén; Los Angeles, CA: 2010.
- Muthén, BO.; Khoo, ST.; Francis, D. *Multistage Analysis of Sequential Developmental Processes To Study Reading Progress: New Methodological Developments Using General Growth Mixture Modeling*. UCLA Center for the Study of Evaluation; National Center for Research on Evaluation, Standards, and Student Testing; Los Angeles, CA: 1998.
- Newman DA. Longitudinal modeling with randomly and systematically missing data: a simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*. 2003; 6(3):328–362. DOI: 10.1177/1094428103254673.
- Newnham EA, Doyle EL, Sng AAH, Hooke GR, Page AC. Improving clinical outcomes in psychiatric care with touch-screen technology. *Psychological Services*. 2012; 9(2):221–223. [PubMed: 22662741]
- Normand ST, Belanger AJ, Eisen S. Graded response model-based item selection for behavior and symptom identification. *Health Services Outcomes Research Methods*. 2006; 6:1–19. DOI: 10.1007/s10742-006-0005-0.
- Patel, V. Cultural issues in measurement and research.. In: Prince, M., et al., editors. *Practical Psychiatric Epidemiology*. Oxford University Press; Oxford: 2003. p. 43-64.
- Patel V. Mental health in low- and middle-income countries. *British Medical Bulletin*. 2007; 81:81–96. DOI: 10.1093/bmb/ldm010. [PubMed: 17470476]
- Patel V, Simunyu E, Gwanzura F, Lewis G, Mann A. The Shona Symptom Questionnaire: the development of an indigenous measure of common mental disorders in Harare. *Acta Psychiatrica Scandinavica*. 1997:469–475. DOI: 10.1111/j.1600-0447.1997.tb10134.x. [PubMed: 9242841]

- Patel V, Araya R, Chatterjee S, Chisholm D, Cohen A, De Silva M, Hosman C, McGuire H, Rojas G, van Ommeren M. Treatment and prevention of mental disorders in low-income and middle-income countries. *The Lancet*. 2007a; 370(9591):991–1005. DOI: 10.1016/S0140-6736(07)61240-9.
- Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. *Lancet*. 2007b; 369:1302–1313. DOI: 10.1016/S0140-6736(07)60368-7. [PubMed: 17434406]
- Prince M, Patel V, Saxena S, Maj M, Maseko J, Phillips MR, Rahman A. No health without mental health. *The Lancet*. 2007; 370(9590):859–877. DOI: 10.1016/S0140-6736(07) 61238-0.
- Psychosocial Working Group. Psychosocial Intervention in Complex Emergencies: A Conceptual Framework. Centre for International Health Studies, Queen Margaret University College; Edinburgh: 2003.
- Pynoos RS, Nader K. Psychological first aid and treatment approach to children exposed to community violence: research implications. *Journal of Traumatic Stress*. 1988; 1(4):445–473. DOI: 10.1002/jts.2490010406.
- Saraceno B, van Ommeren M, Batniji R, Cohen A, Gureje O, Mahoney J, Sridhar D, Underhill C. Barriers to improvement of mental health services in low-income and middle-income countries. *The Lancet*. 2007; 370(9593):1164–1174. DOI: 10.1016/S0140-6736(07)61263-X.
- Saxena S, Thornicroft G, Knapp M, Whiteford H. Resources for mental health: scarcity, inequity, and inefficiency. *The Lancet*. 2007; 370(9590):878–889. DOI: 10.1016/S0140-6736(07)61239-2.
- Smith, CP. Motivation and Personality: Handbook of Thematic Content Analysis. Cambridge University Press; Cambridge: 1992.
- Snyder, C.; McDermott, D.; Cook, W.; Rapoff, M. Hope for the Journey – Helping Children through Good Times and Bad. Westview Press; Boulder, CO: 1997.
- Srinivasan TN, Suresh TR. Non-specific symptoms and screening of non-psychotic morbidity in primary care. *Indian Journal of Psychiatry*. 1990; 32(1):77–82. [PubMed: 21927432]
- StataCorp. Stata Statistical Software: Release 11. StataCorp LP.; College Station, TX: 2009.
- UNICEF. UNICEF; New York: 2009. Machel Study 10-Year Strategic Review: Children and Conflict in a Changing World.
- United Nations Millenium Project. Toward Universal Primary Education: Investments, Incentives, And Institutions, Task Force on Education and Gender Equality. United Nations; New York: 2005.
- Van Ommeren M. Validity issues in transcultural epidemiology. *The British Journal of Psychiatry*. 2003; 182(5):376–378. DOI: 10.1192/bjp.182.5.376. [PubMed: 12724237]
- Veenema TG, Schroeder-Bruce K. The aftermath of violence: children, disaster and posttraumatic stress disorder. *Journal of Pediatric Health Care*. 2002; 16:235–244. [PubMed: 12226592]
- Wang X, Gao L, Shinfuku N, Zhang H, Zhao C, Shen Y. Longitudinal study of earthquake-related PTSD in a randomly selected community sample in North China. *American Journal of Psychiatry*. 2000; 157:1260–1266. DOI: 10.1176/appi.ajp.157.8.1260. [PubMed: 10910788]
- Wilamowska ZA, Thompson-Hollands J, Fairholme CP, Ellard KK, Farchione TJ, Barlow DH. Conceptual background, development, and preliminary data from the unified protocol for transdiagnostic treatment of emotional disorders. *Depression and Anxiety*. 2010; 27(10):882–890. DOI: 10.1002/da.20735. [PubMed: 20886609]
- World Health Organization. WHO DAS 2. World Health Organization; Geneva: 2010.
- Yang H, Stanton B, Cottrel L, Kaljee L, Galbraith J, Li X, Cole M, Harris C, Wu Y. Parental awareness of adolescent risk involvement: implications of overestimates and underestimates. *Journal of Adolescent Health*. 2006; 39(3):353–361. DOI: 10.1016/j.jadohealth.2005.12.008. [PubMed: 16919796]
- Yu X, Lau JTF, Zhang J, Mak WWS, Choi KC, Lui WWS, Zhang J, Chan EYY. Posttraumatic growth and reduced suicidal ideation among adolescents at month 1 after the Sichuan earthquake. *Journal of Affective Disorders*. 2010; 123:327–331. DOI: 10.1016/j.jad.2009.09.019. [PubMed: 19910053]

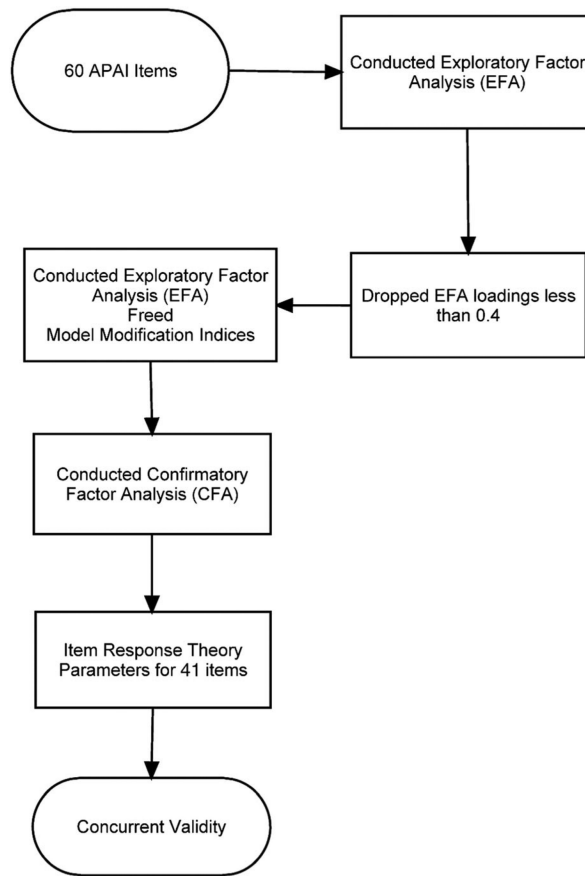


Figure 1. General workflow of the IRT analysis to produce the reduced AYPA measure.

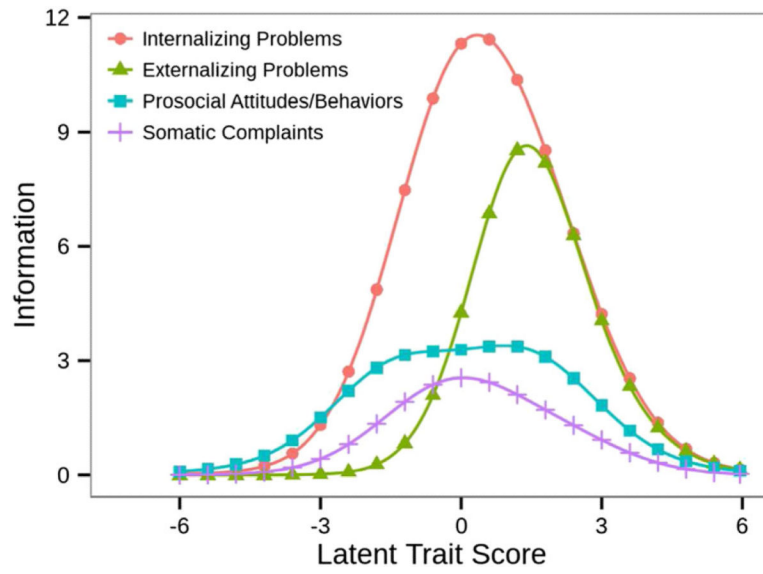


Figure 2.
Information curves for the AYPAs subscales.

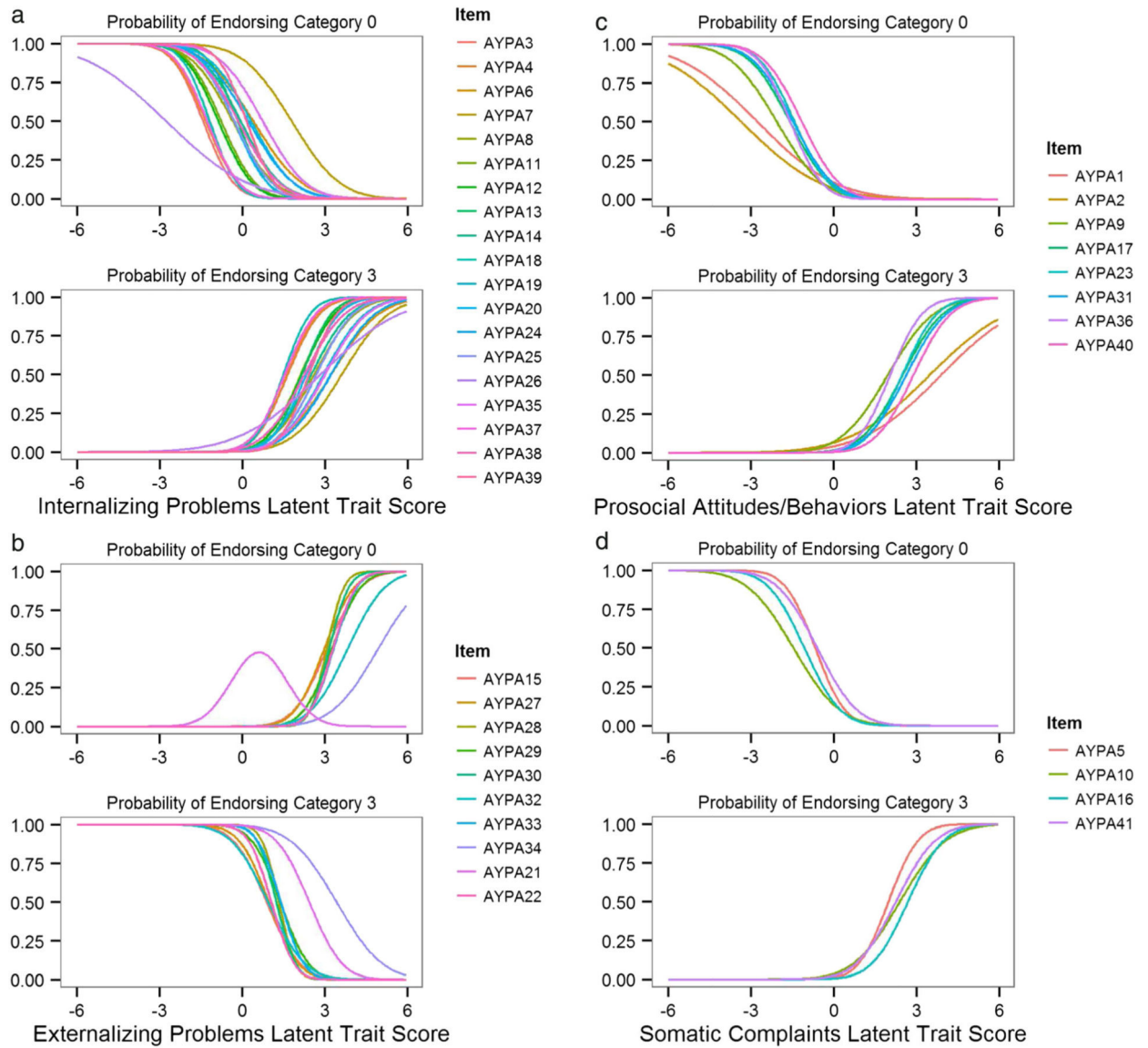


Figure 3. Item characteristic curves for the AYPAs subscales: (a) internalizing subscale; (b) externalizing subscale; (c) prosocial subscale; (d) somatic subscale.

Table 1

Demographic characteristics of Acholi Luo youth study sample

Characteristic	Mean (<i>M</i>) or <i>n</i>	Standard deviation (SD) or (%)	Observed range
Total [<i>n</i> (%)]	667	(100)	
Age [<i>M</i> (SD)]	14.9	(1.1)	[14.0–17.0]
Sex [<i>n</i> (%)]			
Male	298	(44.7)	
Female	369	(55.3)	
Days attend school weekly [<i>n</i> (%)]			
0	187	(28.0)	
2	7	(1.0)	
3	13	(1.9)	
4	28	(4.2)	
5	424	(63.6)	
6	7	(1.0)	
Missing	1	(0.1)	
Highest level of education	5.1	(1.45)	[0.0–9.0]
Physical functioning [<i>M</i> (SD)]	15.3	(8.1)	[0.0–42.0]

Table 2Item summary frequency of full APAI questions ($N = 667$)

	Items	Pe Never	Manok Somewhat	Madwong Often	Tutwal All the time	Missing
APAI 1	I listen to others	87	397	145	38	
APAI 2	I play together with others	60	301	247	55	4
APAI 3	I have a lot of thoughts	94	171	326	76	
APAI 4	I have constant worries	110	172	307	78	
APAI 5	I have pain all over my body	202	215	205	45	
APAI 6	My brain is not functioning well	368	171	106	22	
APAI 7	I think I am of no use	392	125	127	21	2
APAI 8	I think about suicide	568	49	38	12	
APAI 9	I talk constantly about my problems	110	184	317	55	1
APAI 10	I sit alone	274	208	153	32	
APAI 11	I share with others	72	330	184	81	
APAI 12	I lose interest in school	456	57	86	35	33
APAI 13	I get headaches	126	216	274	51	
APAI 14	I lose my appetite	277	244	118	28	
APAI 15	I feel a lot of pain in my heart	193	216	214	44	
APAI 16	I sit with my cheek in my palm	177	223	225	42	
APAI 17	I cry when I am alone	298	166	169	34	
APAI 18	I do not sleep at night	322	179	132	34	
APAI 19	I am disobedient	492	129	37	9	
APAI 20	I feel cold	152	266	228	21	
APAI 21	I share food and eat with others	101	426	105	35	
APAI 22	I lie down all the time (during the day)	425	164	66	12	
APAI 23	I have lots of worries	116	166	302	83	
APAI 24	I want to be alone	386	164	96	21	
APAI 25	I am easily annoyed	254	198	171	43	1
APAI 26	I hold my head	276	204	155	32	
APAI 27	I lose concentration in class	334	103	81	28	121
APAI 28	I drink alcohol	641	23	3	0	
APAI 29	I insult my friends	540	107	17	2	1
APAI 30	I don't greet people	469	80	102	16	
APAI 31	I help others	106	439	95	27	
APAI 32	I don't think straight	368	183	99	17	
APAI 33	I mutter to myself	280	207	157	23	
APAI 34	I don't trust	374	137	125	31	
APAI 35	I feel I can do nothing to help myself	92	122	364	89	
APAI 36	I fight	512	123	24	8	
APAI 37	I use bad language	589	64	12	2	
APAI 38	I am disrespectful	571	55	37	4	
APAI 39	I misbehave	573	73	19	2	

	Items	Pe <i>Never</i>	Manok <i>Somewhat</i>	Madwong <i>Often</i>	Tutwal <i>All the time</i>	Missing
APAI 40	I am disinterested	404	125	112	25	1
APAI 41	I welcome others	112	451	75	29	
APAI 42	I deceive	494	135	35	3	
APAI 43	I am a rough person	585	62	18	2	
APAI 44	I use drugs (like jayi, marijuana)	653	11	1	1	1
APAI 45	I cling to elders	443	116	96	12	
APAI 46	I think I do not have a future	348	123	161	31	4
APAI 47	I am constantly running around	441	128	92	6	
APAI 48	I don't like noise	211	149	236	71	
APAI 49	I think people are chasing me	452	102	93	20	
APAI 50	I have a fast heart rate	231	216	199	21	
APAI 51	I cooperate with others	83	391	148	45	
APAI 52	I fear being alone	227	142	245	53	
APAI 53	I feel sad	115	157	308	87	
APAI 54	I think of bad things	307	151	160	49	
APAI 55	I am weak	230	176	214	47	
APAI 56	I don't feel like talking	368	166	113	20	
APAI 57	I am forgetful	166	235	225	41	
APAI 58	I cry continuously	363	160	123	21	
APAI 59	I do not care whether I live or die	413	111	115	28	
APAI 60	I respect others	145	459	46	17	

Table 3Exploratory factor analysis model fit statistics summary ($N = 667$)

Model fit statistics	3-Factor model	4-Factor model	5-Factor model
χ^2	739	601	534
Degrees of freedom	272	273	268
p -Value	<.001	<.001	<.001
Comparative fit index (CFI)	0.907	0.934	0.947
Root mean squared error of approximation (RMSEA)	0.051	0.042	0.039

Table 4

Confirmatory factor analysis results for four-factor model with factor loadings

	Factor loadings	Standard error
<i>Internalizing</i>		
I have a lot of thoughts (AYPA3)	0.75	0.02
I worry constantly (AYPA4)	0.73	0.02
I think I am of no use (AYPA6)	0.58	0.03
I think about suicide (AYPA7)	0.59	0.05
I sit alone (AYPA8)	0.64	0.03
I feel a lot of pain in my heart (AYPA11)	0.72	0.02
I sit with my cheek in my palm (AYPA12)	0.72	0.02
I cry when I am alone (AYPA13)	0.76	0.02
I do not sleep at night (AYPA14)	0.66	0.03
I have lots of worries (AYPA18)	0.78	0.02
I want to be alone (AYPA19)	0.64	0.03
I hold my head (AYPA20)	0.72	0.02
I don't think straight (AYPA24)	0.61	0.03
I mutter to myself (AYPA25)	0.68	0.02
I think people are chasing me (AYPA35)	0.40	0.04
I feel sad (AYPA37)	0.63	0.03
I think of bad things (AYPA38)	0.74	0.02
I cry continuously (AYPA39)	0.65	0.03
I feel I can do nothing to help myself (AYPA26)	0.40	0.04
<i>Externalizing</i>		
I am disobedient (AYPA15)	0.72	0.03
I drink alcohol (AYPA21)	0.73	0.08
I insult my friends (AYPA22)	0.84	0.03
I fight (AYPA27)	0.76	0.03
I use bad language (AYPA28)	0.90	0.02
I am disrespectful (AYPA29)	0.77	0.04
I misbehave (AYPA30)	0.87	0.03
I deceive (AYPA32)	0.68	0.04
I am a rough person (AYPA33)	0.83	0.03
I use drugs (like jay, marijuana) (AYPA34)	0.61	0.10
<i>Prosocial</i>		
I listen to others (AYPA1)	0.41	0.04
I play together with others (AYPA2)	0.40	0.04
I share with others (AYPA9)	0.60	0.03
I share food and eat with others (AYPA17)	0.65	0.03
I help others (AYPA23)	0.71	0.03
I welcome others (AYPA31)	0.65	0.03
I cooperate with others (AYPA36)	0.73	0.02

	Factor loadings	Standard error
I respect others (AYPA40)	0.68	0.03
<i>Somatic complaints without medical cause</i>		
I have pain all over my body (AYPA5)	0.76	0.03
I get headaches (AYPA10)	0.60	0.03
I feel cold (AYPA16)	0.70	0.03
I am weak (AYPA 41)	0.66	0.03

Table 5Item response theory parameters (a = discrimination) (b = thresholds) for the AYPAs items (N = 667)

	Discrimination	First threshold	Second threshold	Third threshold
<i>Internalizing (total information=50.64)</i>				
I have a lot of thoughts (AYPA3)	0.75	-1.08	-0.26	1.21
I worry constantly (AYPA4)	0.73	-0.97	-0.20	1.19
I think I am of no use (AYPA6)	0.58	0.23	0.76	1.86
I think about suicide (AYPA7)	0.59	1.04	1.44	2.10
I sit alone (AYPA8)	0.64	-0.23	0.59	1.67
I feel a lot of pain in my heart (AYPA11)	0.72	-0.56	0.29	1.51
I sit with my cheek in my palm (AYPA12)	0.72	-0.63	0.25	1.53
I cry when I am alone (AYPA13)	0.76	-0.13	0.51	1.64
I do not sleep at night (AYPA14)	0.66	-0.04	0.68	1.64
I have lots of worries (AYPA18)	0.78	-0.94	-0.20	1.15
I want to be alone (AYPA19)	0.64	0.20	0.93	1.86
I hold my head (AYPA20)	0.72	-0.22	0.58	1.67
I don't think straight (AYPA24)	0.61	0.13	0.94	1.95
I mutter to myself (AYPA25)	0.68	-0.20	0.62	1.82
I think people are chasing me (AYPA35)	0.63	0.46	0.96	1.88
I feel sad (AYPA37)	0.74	-0.95	-0.23	1.12
I think of bad things (AYPA38)	0.65	-0.10	0.49	1.45
I cry continuously (AYPA39)	0.77	0.11	0.79	1.86
I feel I can do nothing to help myself (AYPA26)	0.40	-1.09	-0.47	1.11
<i>Externalizing (total information =27.63)</i>				
I am disobedient (AYPA15)	0.72	0.64	1.48	2.21
I drink alcohol (AYPA21)	0.76	1.76	2.61	0.00
I insult my friends (AYPA22)	0.90	0.88	1.90	2.75
I fight (AYPA27)	0.77	0.73	1.67	2.26
I use bad language (AYPA28)	0.87	1.19	2.03	2.75
I am disrespectful (AYPA29)	0.69	1.06	1.54	2.51
I misbehave (AYPA30)	0.83	0.65	1.58	2.61
I deceive (AYPA32)	0.60	0.65	1.58	2.61
I am a rough person (AYPA33)	0.73	1.16	1.88	2.75
I use drugs (like jayi, marijuana) (AYPA34)	0.84	2.06	2.75	2.97
<i>Prosocial (total information=20.88)</i>				
I listen to others (AYPA1)	0.41	-1.12	0.60	1.58
I play together with others (AYPA2)	0.40	-1.34	0.11	1.39
I share with others (AYPA9)	0.60	-1.24	0.26	1.17
I share food and eat with others (AYPA17)	0.65	-1.03	0.81	1.62
I help others (AYPA23)	0.71	-1.00	0.90	1.75
I welcome others (AYPA31)	0.65	-0.96	1.01	21.71
I cooperate with others (AYPA36)	0.73	-1.15	0.56	1.50

	Discrimination	First threshold	Second threshold	Third threshold
I respect others (AYPA40)	0.68	-0.78	1.31	1.95
<i>Somatic complaints without medical cause (total information = 11.63)</i>				
I have pain all over my body (AYPA5)	0.76	-0.52	0.32	1.50
I get headaches (AYPA10)	0.60	-0.88	0.03	1.43
I feel cold (AYPA16)	0.70	-0.75	0.32	1.86
I am weak (AYPA 41)	0.66	-0.40	0.28	1.47