

UCLA

UCLA Previously Published Works

Title

Development of the South Sudan Mental Health Assessment Scale

Permalink

<https://escholarship.org/uc/item/25w3x3g2>

Journal

Transcultural Psychiatry, 59(3)

ISSN

1363-4615

Authors

Ng, Lauren C
Solomon, Jordan S
Ameresekere, Maithri
[et al.](#)

Publication Date

2022-06-01

DOI

10.1177/13634615211059711

Peer reviewed

Development of the South Sudan Mental Health Assessment Scale

*Lauren C. Ng¹, Jordan S. Solomon², Maithri Ameresekere³, Judith Bass⁴, David C. Henderson⁵, and Shubha Chakravarty²

¹University of California, Los Angeles

Boston University School of Medicine and Boston Medical Center

²The World Bank

³University of California, San Francisco

⁴Johns Hopkins Bloomberg School of Public Health

⁵Boston University School of Medicine and Boston Medical Center

***Corresponding author:**

Lauren C. Ng,

Department of Psychology, University of California, Los Angeles,

1285 Franz HallBox 951563, Los Angeles, CA 90095-1563, USA.

Email: laurenng@ucla.edu

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was made possible through the financial support of the World Bank's multi-donor trust fund for adolescent girls, the Umbrella Facility for Gender Equality, and the Bank-Netherlands Partnership Program. LN's time was funded by NIMH (T32 MH116140; K23MH110601).

Abstract

This study developed and validated a measure that captures variation in common local idioms of distress and mental health problems experienced by women in South Sudan, a country which has experienced over 50 years of violence, displacement, and political, social, and economic insecurity. This measure was developed during a randomized controlled trial of the Adolescent Girls Initiative (AGI) and used qualitative Free Listing (n=102) and Key Informant interviews (n = 27). Internal reliability and convergent validity were assessed using data from 3,137 randomly selected women (ages 14–47) in 100 communities in South Sudan. Test-retest and inter-rater reliability were assessed using responses from 180 women (ages 15–58) who completed the measure once, and 129 of whom repeated the measure an average of 12 days (SD=8.3) later. Concurrent validity was assessed through the ratings of 22 AGI leaders about the presence or absence of mental health symptoms in the 180 women in the test-retest sample. The study resulted in the development of the South Sudan Mental Health Assessment Scale, a 24-item measure assessing six idioms of distress. The scale consisted of one factor and had excellent internal, test-retest, and interrater reliability. The scale also demonstrated good convergent and concurrent validity and performed well psychometrically. Moreover, its development provides an example for other organizations, working in environments where mental health measures have not yet been developed and validated, to create and validate measures relevant to their populations. In this way, the role of mental health in development settings can be more rapidly assessed.

Keywords: South Sudan; mental health; psychometrics; trauma; validation.

Introduction

Globally, mental and substance use disorders are the leading cause of years lost to disability (Whiteford et al., 2013). Therefore, population mental health may be a critical factor in the success of poverty alleviation interventions, particularly in fragile or conflict-affected settings where substantial trauma has occurred, and where poverty and poor mental health may adversely impact each other. However, there are often few, if any, measures that have been developed and validated in conflict-affected settings. This dearth of validated measures is highly problematic since the way that people experience, express, and describe mental health problems can vary by culture, context, and language (Hinton & Lewis-Fernández, 2010).

This study developed and validated a measure of mental health symptoms in South Sudan, a country that has been exposed to over 50 years of conflict and has very little development, ranking 169th out of 188 countries on the UN Human Development Index (United Nations Development Programme, 2018). Following decades of war, South Sudan gained independence from Sudan in 2011, becoming the world's youngest country to date. Unfortunately, civil war broke out in South Sudan in December 2013 and has continued intermittently through the beginning of 2020. During this time, hundreds of thousands of South Sudanese have been killed and millions displaced. Data collection for this study started in November 2013 and concluded in April 2015, during the midst of the beginning of the most recent years-long conflict.

As a result of decades of violence and war, the majority of the population has been exposed to high rates of violence, displacement, and political, social, and economic insecurity, with documented high rates of mental health problems like depression and post-traumatic stress disorder (PTSD) (Ayazi et al., 2012; Friedrich, 2016; Ng et al., 2017). Rates of depression have

been estimated to be approximately 16% (Ayazi et al., 2012), while PTSD prevalence rates in South Sudan have been estimated to range from 36% to 48% (Ayazi et al., 2012; Karunakara et al., 2004; Ng et al., 2017; Roberts et al., 2009). Moreover, the conflict and subsequent underdevelopment have resulted in almost 90% of people in South Sudan facing severe deprivation across multiple domains including schooling, nutrition, sanitation, electricity, cooking oil, and wealth (Ballon & Duclos, 2015). Low literacy and lack of income-generating activities affect almost all in South Sudan, but are particularly difficult for young women (Ballon & Duclos, 2015). In South Sudan, women are almost twice as likely to be illiterate as men, and only about half of all youth are employed (Ballon & Duclos, 2015). Lack of education and income increase vulnerability to early marriage and childbearing, and in South Sudan almost 45% of women are married before the age of 18 and a third have had children by then (Ministry of Health & National Bureau of Statistics, 2013).

In order to improve employment and income outcomes and break the cycle of early marriage and childbearing for adolescent girls and young women in South Sudan, the Adolescent Girls Initiative (AGI) (World Bank, 2016) was implemented by BRAC (originally known as the Bangladesh Rural Advancement Committee, but since 2009, going only by the abbreviated name) and financed through the World Bank's multi-donor trust fund on adolescent girls. The South Sudan AGI program was adapted from BRAC's Empowerment and Livelihoods for Adolescents (ELA) which was originally developed in Bangladesh and previously implemented in Uganda and Tanzania (Bandiera et al., 2015). The program provides livelihood and life skills training to young women aged 15 to 24. The program was implemented in 100 communities in five districts in South Sudan to improve the life skills and livelihoods of 3,000 vulnerable

adolescent girls. This study was conducted as a component of a larger Randomized Controlled Trial (RCT) of the AGI in South Sudan.

Given the extensive and widespread trauma exposure in South Sudan, the young women enrolled in the AGI intervention faced unique life circumstances that were different from those who had participated in Uganda, Tanzania, and Bangladesh. Therefore, the World Bank researchers hypothesized not only that mental health issues such as depression, anxiety, and posttraumatic stress may be widespread amongst participants, but also that participants with mental health concerns may not participate in the AGI program and/or may not benefit from it in the same way as participants who do not have mental health concerns. Mental health was hypothesized to be both a potential outcome of the program and a possible mediator or moderator of the program's impact on other outcomes, such as employment. In order to capture the potential effects of mental health on the AGI intervention, the researchers conducting the RCT on the impacts of the AGI program in South Sudan sought to include a locally validated measure of psychological well-being symptoms that might be amenable to change during an economic strengthening intervention. Although some studies have assessed or treated mental health in South Sudan and with South Sudanese refugees (Adaku et al., 2016; Ayazi et al., 2013; Badri et al., 2012; Lien et al., 2016; Meffert et al., 2010; Roberts et al., 2009; Schweitzer et al., 2006; Tol et al., 2018), no mental health assessments could be found that had been validated for use in South Sudan. Therefore, the researchers developed their own mental health assessment, using the DIME Program Research model (Center for Refugee & Disaster Response, 2016). The DIME program was selected because it provided detailed manualized instructions for developing and validating mental health assessment tools based on local idioms of distress. The instructions could be carried out by non-mental health professionals, and in fact specifically targeted non-

governmental organizations (Applied Mental Health Research Group, 2013). The DIME approach begins with qualitative assessment in order to understand how people in the population of interest describe, prioritize, perceive, and cope with mental health and psychosocial problems (Applied Mental Health Research Group, 2013). The DIME approach then uses the results of the qualitative assessment to draft, translate, and validate the instruments (Applied Mental Health Research Group, 2013). This study describes the development and validation of this measure, the SSMHAS, using modules one and two of the DIME approach.

Methods

The AGI Intervention

AGI girls' clubs were established in five urban and peri-urban areas in South Sudan: Juba, Torit, Yei, Rumbek, and Bor. These community-level girls' clubs met five times per week for two years with a trained female adolescent club leader. Activities included sports and games, life-skills training including sexual and reproductive health, and financial literacy training. In addition, the program financed the young women's participation in short livelihood training courses in one of 10 trades. Although the primary objective of the program was to strengthen the livelihood activities of participants, the program was conceived to holistically address multiple challenges in young women's lives, specifically reducing social isolation and building social support structures to improve well-being.

Participant selection

The BRAC team in South Sudan selected 120 clusters, or communities, to be part of the RCT of the AGI clubs. In each of five areas, 12 clusters were randomly selected to receive the intervention and 12 clusters were randomly selected to be in the control group, which did not receive any intervention. The clubs were established between September and December 2010

and the program finished in June 2013. The clubs were open to all girls and women aged 15–24 who lived in the villages that received the AGI. Each club formally had 30 club members and one club leader, but in practice additional girls and women participated in club activities.

Interviewers

All qualitative and quantitative interviews were conducted by six South Sudanese interviewers, each having at least some post-secondary education and having demonstrated high-quality work in previous research engagements with BRAC. All interviewers were South Sudanese women aged 18–30, with varied ethnicity, similar to the target population of the study, and were fluent in Juba Arabic and English.

Interviewers participated in a four-day classroom-based training on mental health, research ethics, and research methods for Free Listing. The training included background information such as definitions of mental health, and a discussion of why US or European mental health instruments may not be relevant in South Sudan. Trainees also learned proper research protocol such as informed consent, non-judgment, asking non-leading questions, and probing for additional information. In addition, interviewers practiced conducting informed consent and the interviews, and were tested with potential difficult situations they may face during interviews.

Ethical approval and consent

All respondents were informed that participation was voluntary and, due to low literacy rates, all respondents gave verbal informed consent. Informed consent from the household head was also sought for the $n = 3,137$ person baseline survey before participants were approached about study participation. For the qualitative and reliability studies, permission of the local leader including the chief, headmaster of boarding schools, or village elders was sought before participants were approached for study participation. Parental consent for minor participants was

not sought for the qualitative and reliability studies because many of the participants were internally displaced persons or in boarding schools, so parental permission would have been infeasible. In line with World Bank protocol and due to the active ongoing conflict in South Sudan, IRB approval was not sought for the data collection. However, the World Bank received permission to conduct research from the South Sudanese Ministry of Finance and Economic Planning and the Ministry of Gender, Child and Social Welfare. The research team also received permission from the Kator and Munuki Payam offices in Juba and local leaders (headmaster, market leader, chief) to interview members of their community. Approval for data analysis of the deidentified collected data was obtained from the Partners HealthCare Institutional Review Board.

Remuneration

Participants of the study were given soap and/or sanitary pads as an appreciation for taking the time to participate in the study. Participants were given the remuneration after the interviews were complete and were not told about the remuneration before the study. The market value of the remuneration was approximately US\$1–3.

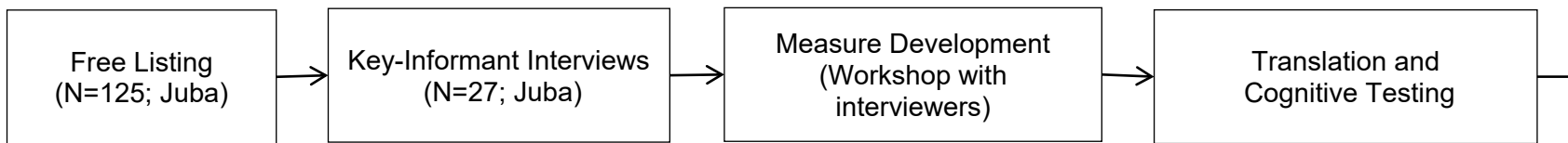
Phase 1: Qualitative assessment to identify and describe mental health problems

The development and evaluation of the SSMHAS followed the first three modules of the DIME Program Research Model including conducting a qualitative assessment, developing draft instruments, and validating instruments (Center for Refugee & Disaster Response, 2016). See Figure 1 for a flowchart of all study activities.

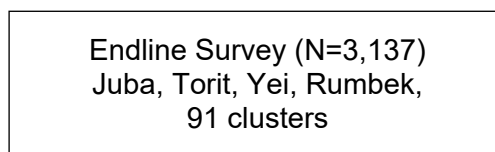
Figure 1. Flowchart of study activities

AGI Intervention; communities in Juba, Torit, Yei, Rumbek, Bor

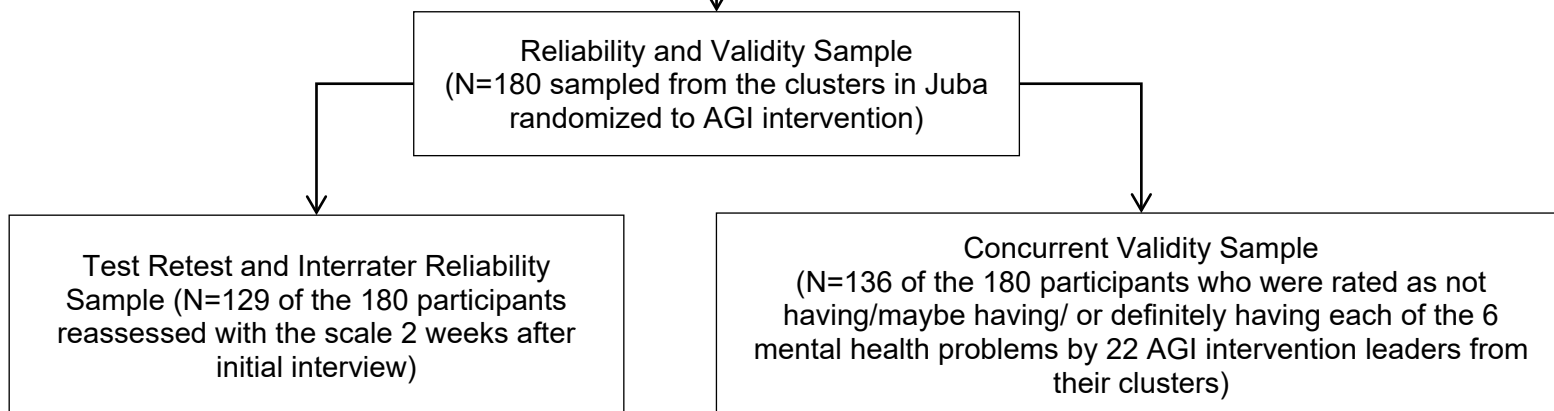
Phases 1 & 2: Qualitative assessment and development of instrument



Phases 3 & 4: Endline survey data used to assess dimensionality, internal reliability, and convergent validity



Phase 3: Reliability and validity study



Free Listing exercise

Mental health problems and local idioms of distress (Hinton & Lewis-Fernández, 2010) were identified through Free Listing interviews in Juba between November 2013 and October 2014. The first rounds of Free Listing enrolled women aged 18 to 28 and began in November 2013, but were placed on hold when conflict broke out in December 2013. The second round of Free Listing took place in Juba in October 2014 with a more diverse group of women. In total, data were collected with 125 women aged 14–48.

Participants

The team hoped to identify respondents who came from varied backgrounds, thus respondents were chosen using the “Maximum Variation Sampling” technique (Center for Refugee & Disaster Response, 2016). The research team selected a number of key variables on which they wanted respondents to vary, including schooling (high-quality schools, low-quality schools), SES (wealthy neighborhoods, middle-class neighborhoods, poor neighborhoods), exposure to the 2013 conflict (the interviewers identified neighborhoods in Juba with high, medium, and low conflict exposure), and employment (young women in school, women who were working in shops and the market, and women who did not work outside of the home). The interviewers identified schools, communities, and marketplaces that would have women aged 18–28 from each of these categories. The research team then contacted the head of each community (i.e. the school's headmaster, the local leader in the market, or the local chief of a neighborhood) to select women, primarily between the ages of 18–25, who would be interested in participating in the study. For logistical reasons, the interview teams spent one day at each interview site. All interviews were voluntary and verbal informed consent was obtained from each respondent.

In total, 54 Free Listing interviews were conducted with a total of 102 participants (ages 14–48; $M[SD] = 21.6[7.2]$). Of the 54 interviews, 48 included two respondents while six included only one respondent. Interviews were conducted in 11 communities in Juba that varied in terms of exposure to the 2013 conflict and SES and were similar to the RCT sites in Juba. Key Informants included parents, siblings, chiefs, police officers, headteachers, and teachers. In total, 27 Key Informants were identified (ages 18–60, $M[SD] = 36.3[13.2]$).

Procedure

The Free Listing exercise consisted of pairs of interviewers asking open-ended questions to pairs of respondents (in six interviews there was just one respondent). Respondent pairs were used because BRAC staff previously observed that respondents were much more comfortable providing answers when paired with another respondent. Respondents would often bounce ideas off other respondents and create a dialogue that elicited a higher frequency and a more honest set of responses. Typically, one interviewer would ask the questions while the other interviewer recorded the responses. If one interviewer got tired of writing they were free to switch tasks.

Participants responded to the following questions: 1) “What are the problems that young women like you have in this community?” 2) “What are the thoughts, feelings, and behaviors that young women have related to these problems?” 3) “How would you describe a person who had these problems but is now doing well?” and 4) “Who are the people that young women go to when they have these problems?” The respondents were instructed to talk about problems that young women faced in the community and were encouraged not to talk about themselves. If they were speaking in the first person, they were asked if this was also a problem faced by many in the community. The respondents were instructed to list a “problem” and provide a short description of it, each of which interviewers wrote verbatim. If the respondent told a long story

or provided a long description, the interviewer would ask them to summarize. Once the respondents finished listing their problems, the interviewers probed in a non-leading way to see if there were any more problems.

At the end of the Free Listing exercises, the interviewer asked respondents to describe whom young women go to when they have problems. Respondents provided the contact information of specific people in their communities, which was used to generate the list of participants for the Key Informant interviews.

Data analyses

The Free Listing results were analyzed by all of the interviewers during a workshop. At the workshop the interviewers were split into three teams, with each assigned one of the questions from the Free Listing exercise. The teams then made a master list of all the: 1) Problems, 2) Thoughts, feelings and behaviors, and 3) Ways you can tell a young woman is doing well. The interviewers counted how often each problem was mentioned. Following procedures outlined in the DIME manual (Center for Refugee & Disaster Response, 2016), once each team had a full count of all the lists, the interviewers reviewed, discussed, and provided examples of each problem as a group. The interviewers were encouraged to think about the causes of each problem, how the problem manifests itself in the behavior of young women, how the problem affects people in their community, and the frequency of this problem in their communities. After a definition of each problem from the Free Listing exercise, they were grouped into overarching problems.

Results

The Free Listing exercise generated 49 different mental health problems experienced by young women in South Sudan. For each of the 49 problems, the interviewers decided if the

problem was closely related to a previously mentioned problem or was a distinct problem, and grouped the problems that were closely related into one construct. The interviewers then chose one word that best describes and summarizes the problem. Eleven problems were identified and named.

Key Informant interviews

Using the problem list generated from the Free Listing interview, a questionnaire was developed to learn more about these 11 problems from Key Informants who were identified during the Free Listing exercise.

Participants

Key Informant interviews were conducted with 27 participants who were identified by Free Listing participants as people in their community that young women and girls turn to when they have problems. Identified Key Informants were contacted and asked to participate in qualitative interviews. All respondents were informed that participation was voluntary and all respondents gave verbal informed consent.

Procedure

Interviewers described the Free Listing exercises and presented the Key Informants with a list of the 11 most common “problems” identified in the Free Listing exercises. Key Informants were encouraged to choose problems that they were familiar with to discuss in greater detail. The following open-ended questions were asked about each problem: (1) “What are the causes of this problem?” (2) “What have young women told you about this problem?” (3) “What are the effects or reactions to this problem? How does it affect the community?” (4) “What are some solutions to this problem?” and finally, space was provided in case respondents wanted to tell additional stories about their experiences with this problem in the community. Respondents were

encouraged to select and discuss many of the 11 problems but typically only chose to discuss one or two problems.

Data analyses

Interviewers reviewed the Key Informant interview notes and identified ways in which each problem was described, using the exact words of the interviewees. If wording was identified that described an already identified construct or symptom, but did so with new wording that was more understandable and superior, then the old wording was replaced with the new wording.

Free Listing and Key Informant results

Results from the Key Informant interviews resulted in the 11 most common mental health problems being identified and described. Of these 11, seven problems were selected for inclusion in the SSMHAS: Pain in heart (waja gelba), Stressed up (nefsiat), Shame (fadia), Rudeness (takian), Isolated or abandoned (abau/hasud), and Traumatized (kafu). See Table 2 for symptoms associated with each mental problem. Abau and hasud were collapsed into one because the descriptions from the Free Listing and Key Informants were very similar, creating six distinct items. The items for bad feelings / unhappy were not included because they were deemed overly broad and part of all of the other constructs, much the same way as negative affect is part of multiple separate mental health constructs such as depression, anxiety, and posttraumatic stress disorder. The remaining two items (crazy and witchcraft) were not included because they were deemed to be associated with severe mental illness symptoms, such as psychosis, which were not hypothesized to be affected by the AGI intervention. See complete results and descriptions in Table 1.

Table 1

Problems Generated from the Free Listing Exercise

Identified thoughts, feelings, or behaviors	# of responses	Illustrative Quotes	Included? If no, for what reason?
Stress up	17	<p>“Stress up is a mental problem for both woman. She want to do something there is no way out even to get solution for that problem there is no way out the problem is continuously.”</p> <p>“Too much thinking without any single rest and no appetite for eating food.”</p>	Yes
Bad feelings	16	<p>“Due to bad feeling sometimes you feel like migrate away from the people and go to the bush and stay with the animals because people are causing him bad feeling.”</p> <p>“I will know that the person has bad feels when the person doesn't talk to people and always unhappy and annoyed.”</p> <p>“She doesn't want people to pick her things and not ready to interact to people She does not like joking, when you trying to joke with her she will wants to fight people.”</p> <p>“She will stay alone and does not want to playing with others. She will only want to talk to people and she does not want people to talk to her. When you started she will not respond.”</p> <p>“If she did some bad she not accept her mistakes.”</p>	Included as symptom of pain in heart
Committing suicide	14	<p>“How will I face people or pass by people at least when you killed yourself it will be better.”</p>	No, beyond the scope of the program.
Annoyed	14	<p>“You will like not talking to anybody and also like to fight with anyone.”</p>	No, overly broad
Unhappy	14	<p>“You will not be happy because there a lot of thing you want to get but no way.”</p>	No, overly broad
Too much thinking	14	<p>“You don't sleep at night thinking all the time about the problem she has.”</p>	Included as symptom stress up
Lonely	14	<p>“She will always stay lonely because parents are refusing her that she is married and husband is not taking care.”</p>	Included as symptom to isolation
Sadness	12	<p>“A person feel like crying always and not happy anymore.”</p>	No, overly broad

Fear	10	“Not staying freely, you will fear be move and take to any one because of which condition you are in.”	Yes, included as symptom to traumatized
Shameful	9	“You think always people will laugh at you over the problem you have.” “Shame is a nature and is some one who feel he has done some thing which is not liked by the society.” “Something which is normally happening and the person do it shame also develops naturally.” “The person will fear to expose herself in the society, fear to talk openly and even fear to see somebody's eyes.” “The person also feels guilty all the time, will not have capacity to lead a group or being a leader.”	Yes
Weakness	9	“This occur always when you have a lot of problem you will lose weight lack of appetite.”	Yes, included as symptom to pain in heart
Thin	9	“Loss of appetite to eat would make someone sick.”	No, not inherently a mental health problem
Quiet	9	“A person is not ready to talk to any one else in the community she belong.”	Yes, included as symptom to shame
Unhappy	8	“Because way thing is destroyed by the current crisis in the country.”	No, overly broad
Rudeness	7	“The face becomes red and always ready to fight anybody.”	Yes
Stress	6	“See Stress Up”	Yes, see stress up
Migrations	5	“They will feel like moving away from the area where there is problem.”	No, outside the scope of the program
Worried	5	“They feel worried because of the things that happened during the war.”	Yes, see stress up
Heart attack	5	“When their problem also you don't feel well also the heart is not settle.”	Yes, included as "pain in heart"
No respect	5	“No listening to anybody at home who tried to talk to you.”	Yes, see rudeness
Isolated	5	“You stay alone and without any one else.”	Yes
Discriminated	5	“When early problems occur you will think that the world come to you alone.”	No, beyond the scope of the program

Traumatize	4	<p>“Someone who has been in the war area witness a lot of things like killing, beating, robbing and being rude to others.”</p> <p>“And also people who have not been war areas they are now fears like the guns shot in Juba so cause someone to be traumatized.”</p> <p>“Some of children carry heavy load on their head and beatings that make them to be traumatized. And this is why some of student want to adopt such system.”</p> <p>“Traumatize always cause psychology problems like instability in the mind that led to craziness.”</p>	Yes, included with fear
Crying	4	“Shedding the tears if such problem occur.”	Yes, included as symptom to pain in heart
Fighting	4	“They always feel like fighting especially if the husband is drinking alcohol.”	Yes, included as symptom to rudeness
Helpless	4	“You feel that there is no one to help you and stay alone.”	Yes, included as symptom to isolated
Madness	4	“Someone moving on the way talking alone to himself.”	No, beyond the scope of the program
Discourage	4	“You feel like this is not my country why all these happen to my states.”	No, related to shame
Silences	4	“A person is not ready to open her mouth to talk to any body.”	Yes, included as symptom of shame

Table 2

SSMHAS Descriptives from the Endline and the Reliability and Validity Samples

How often do you... (%)	Endline (N=3,136)						Reliability and Validity Sample (N=180)							
	Never=1	Rarely=2	Frequently=3	Always=4	Mean	SD	Never	Rarely	Frequently	Always	Mean	SD	t(3314)	p
Have misunderstandings	43.11	48.76	6.47	1.66	1.67	0.67	25.56	68.89	5.56	0	1.80	0.52	2.62	0.009
Have a quick temper	35.24	50.64	12.63	1.5	1.80	0.71	26.11	59.44	13.89	0.56	1.89	0.64	1.58	0.115
Get into fights or quarrels	40.94	46.52	11.8	0.73	1.72	0.69	33.89	61.67	3.89	0.56	1.71	0.56	-0.23	0.818
Feel rude (<i>takian</i>)	41.68	47.23	10.04	1.05	1.70	0.69	40	50.56	9.44	0	1.69	0.64	-0.20	0.845
Have many thoughts	23.5	48.15	19.45	8.9	2.14	0.88	20	58.33	17.22	4.44	2.06	0.74	-1.15	0.252
Have sleepless nights	27.93	53.44	14.8	3.83	1.95	0.76	30	50	17.78	2.22	1.92	0.75	-0.39	0.694
Feel unsettled	45.36	41.28	11.07	2.3	1.70	0.75	52.78	45.56	1.11	0.56	1.49	0.55	-3.66	<.001
Feel stressed up (<i>nefsiat</i>)	36.32	47.42	13.01	3.25	1.83	0.77	30.56	54.44	14.44	0.56	1.85	0.67	0.31	0.759
Feel lonely	36.89	50.61	9.38	3.12	1.79	0.74	54.44	41.67	3.89	0	1.49	0.57	-5.24	<.001
Feel unwanted	55.26	35.33	8.29	1.12	1.55	0.69	78.89	18.33	2.78	0	1.24	0.49	-5.99	<.001
Feel useless	59.12	32.49	6.66	1.72	1.51	0.70	79.44	18.33	1.67	0.56	1.23	0.50	-5.24	<.001
Feel isolated or abandoned (<i>hasud or abau</i>)	50.83	39	9.12	1.05	1.60	0.70	78.89	18.89	2.22	0	1.23	0.47	-7.05	<.001
Feel weak	26.63	62.05	9.6	1.72	1.86	0.64	12.22	82.22	5.56	0	1.93	0.42	1.43	0.154
Feel sorrowful	23.85	61.1	13.33	1.72	1.93	0.66	17.78	78.33	3.89	0	1.86	0.45	-1.37	0.172
Feel like crying	32.94	51.85	13.01	2.2	1.84	0.72	31.11	58.89	10	0	1.79	0.61	-1.01	0.310
Feel pain in heart (<i>waja gelba</i>)	23.92	56.57	16.26	3.25	1.99	0.73	40	37.78	20.56	1.67	1.84	0.81	-2.66	0.008
Are you easily startled	47.8	39.51	11.38	1.31	1.66	0.73	31.11	61.11	7.78	0	1.77	0.58	1.89	0.058

Think about bad things that happened in the past	29.21	54.02	13.55	3.22	1.91	0.74	40	52.22	6.67	1.11	1.69	0.65	-3.88	<.001
Do bad things in the past make you act crazy	50.26	38.3	9.25	2.2	1.63	0.74	91.67	7.78	0.56	0	1.09	0.30	-9.82	<.001
Feel traumatized (<i>kafu</i>)	36.99	48.57	12.6	1.85	1.79	0.73	55.56	38.33	6.11	0	1.51	0.61	-5.21	<.001
Fear that others are talking about you	39.38	47.93	10.55	2.14	1.75	0.72	68.89	27.22	3.33	0.56	1.36	0.58	-7.26	<.001
Feel shy	36.61	50.16	11.58	1.66	1.78	0.71	37.22	57.78	5	0	1.68	0.57	-1.96	0.051
Hide from others	59.28	33.07	6.47	1.18	1.50	0.67	86.11	12.22	1.67	0	1.16	0.41	-6.73	<.001
Feel shame (<i>fadia</i>)	33.16	49.3	14.03	3.51	1.88	0.77	35.56	49.44	12.78	2.22	1.82	0.74	-1.05	0.293
Mean scale score (M, SD)					1.77	0.41					1.63	0.25	-4.60	<.001

Note. Items listed above each of the common mental problems (e.g., *takian*, *nefsia*, etc.) are symptoms of that mental problem, for example feeling weak, feeling sorrowful, feeling like crying are all symptoms of pain in the heart (*waja gelba*)

Each of the six problems were presented to the interviewers. The interviewers discussed the similarity and differences between each problem to ensure that they were distinct and three key symptoms from each problem were identified. A scale was created where each of the six problems had a list of three symptoms, and then the name of the problem itself was included as a symptom, resulting in a 24-item scale. See Table 2 for the SSMHAS item descriptives. Possible responses were provided on a Likert Scale of 1 (never), 2 (rarely), 3 (frequently), and 4 (always). Therefore, the scale score range was 24 to 96. A “sometimes” response was not included based on previous work in South Sudan that found many responses would bunch at “sometimes.”

Phase 2: Developing the SSMHAS

Comparison with existing scales

The Free Listing and Key Informant data was entered into a spreadsheet and shared with MA, JB, and SC. MA and JB conducted a search of general mental health scales which had symptoms that matched the problems from the qualitative exercises. In the end, the research team could not find an existing scale that adequately matched the problems from the qualitative work. JS worked with MA, JB, and the interviewers to develop a scale using the exact wording from the qualitative exercises.

Translation

The endline survey, including the AGI scale, was written in English. Research staff verbally translated each question when interviewing the respondent. This was done because the predominant language in the study area, Juba Arabic, does not have a written form. In addition, the questions were not translated into Arabic because not all research staff read Arabic, though they all read English. During training, research staff discussed and came to consensus on the wording of all questions in Juba Arabic.

Cognitive/pilot testing of the instrument

Interviewers practiced conducting the assessment tool with non-sampled women and girls in Juba. The survey was adjusted based on feedback from the interviewers during this testing period.

Phase 3: Assessing scale reliability and dimensionality

Descriptives, dimensionality, and internal reliability

Participants and procedure

In 2014 and 2015, the World Bank research team administered the SSMHAS to 3,137 adolescent girls and women (ages 14–47, $M = 22.04$, $SD = 4.03$) as part of the endline survey of a randomized control trial of the AGI. The survey was administered to 35 randomly selected girls and young women from each of the intervention and control clusters in the study. Due to insecurity, the endline survey was not conducted in Bor so only 91 clusters from four areas were surveyed. Results from this endline survey data were used to assess the factor structure and internal reliability of the developed measure.

Analysis

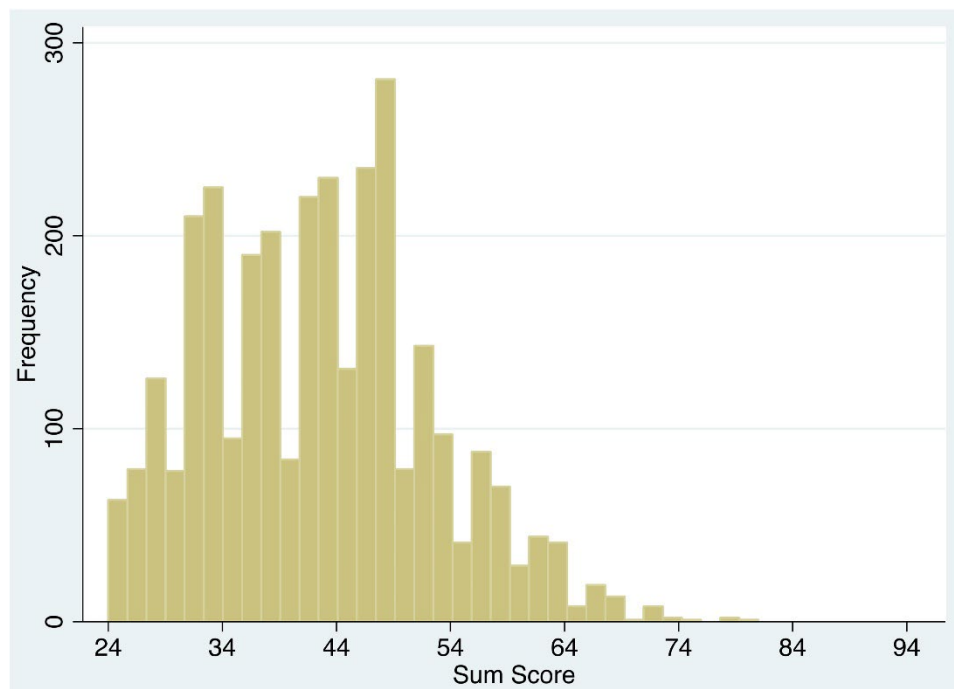
Descriptive statistics of the SSMHAS were calculated at the item level and the scale level using data from the endline survey. Exploratory factor analysis using principal factors was also conducted to assess whether the SSMHAS was measuring a uni- or multi-dimensional underlying mental health construct. The minimum eigenvalue to be retained was set to 1. To assess internal reliability, Cronbach's alpha was run on the complete scale and those identified through the factor analysis.

Results

Results from the endline survey indicated that the mean sum score was 45.50 ($SD = 9.82$) with a range of 24 to 81. Scores on the SSMHAS were somewhat skewed such that three-quarters (75.26%) of endline participants reported that symptoms were “rarely” or “never” experienced, one-quarter (24.55%) reported that symptoms were “rarely” to “frequently” experienced, and less than one percent (0.19%) reported that symptoms were “frequently” to “always” experienced (see Figure 2). On average, scores for individual items and the total scale score fell between “never” and “rarely” (see Table 2 for more detail). The individual items with the highest mean score were “having many thoughts” with a mean of 2.14 (rarely) and a SD of 0.88, followed by “feeling pain in heart” with a mean of 1.99 and an SD of 0.73. As part of the SSMHAS, participants were asked how often they experienced each of the six primary idioms of distress (e.g., “How often do you feel rude (takian)” “How often do you feel pain in heart (waja gelba)?”). While the majority of participants reported few mental health problems on the SSMHAS, 332 (10.59%) reported experiencing at least one of the six primary mental health problems “always” and

1,362 (43.43%) reported experiencing at least one of the six primary mental health problems at least “frequently.”

Figure 2: Distribution of endline scores on the SSMHAS.



Results of the factor analysis of the scale from the endline survey found that one factor fit the data best and explained the 83.12% of the variance (eigenvalue = 7.15; see Supplemental Table 1). The internal reliability of the complete scale was also excellent, with an overall Cronbach's alpha of .91 from the 3,137-person endline sample and .80 from the 180-person validity and reliability sample.

Test-retest and interrater reliability

Participants and procedure

To assess the test-retest and inter-rater reliability of the measure, 180 women (ages 15–58) who participated in the intervention in Juba were administered the mental health scale as part of the reliability and validity sample. The research team worked with AGI club leaders and BRAC community organizers to contact all available club participants in Juba and 178 (98.89%) of the original club participants in Juba were identified and interviewed. All but two interviews were conducted in Juba Arabic (one was conducted in English, and one was conducted in Dinka). All participants were re-contacted, and an

average of 12.40 days later (SD = 8.29; range = 0 to 36 days) the 129 who were available and agreed to be re-interviewed were administered the measure again. Of the 129, 51 (39.84%) were interviewed by the same interviewer and 77 (60.16%) were interviewed by a different interviewer (the name of the first interviewer was missing for one participant). Participants were randomly assigned to receive the same or a different rater. In total, four interviewers participated in this portion of the study. Interviews were conducted between February and April 2015. All respondents were informed that participation was voluntary and all respondents gave informed consent.

Analyses

A repeated measures t-test was used to assess whether scale scores at time one varied significantly from those at time two. To assess the size of the discrepancy between the test and the retest scores, absolute values of the differences of the scores were also calculated. Test-retest reliability was assessed using Intraclass Correlation Coefficient (ICC) analyses. Interrater reliability was assessed by computing ICC on the subsamples of participants who were rated at both time points by the same rater versus a different rater. To assess whether the sizes of the discrepancy between test and retest scores differed when assessments were conducted by the same or different rater, an independent-sample t-test was also conducted.

Results

Results from a paired t-test of the scale scores of 129 women of the original 180 who were re-administered the measure found that there were no significant differences in scale scores from time one to time two (M(SD) from T1 = 38.91 (5.96), M(SD) from T2 = 39.21 (5.89); $t(128) = -0.48, p = .63$). In addition, the absolute value of the size of the discrepancy between mean scores on the test and the retest was 5.41 (SD = 4.43), with more than 50% of participants having test and retest scores within four points of each other on a scale that ranged from 24 to 96.

Since the item-level data were skewed, for the test-retest reliability analysis item-level scores were recoded into a binary scale, such that scores of 1 (never) and 2 (rarely) were recorded as 0, and scores of 3 (frequently) and 4 (always) were recorded as 1 and ICC was computed at the item level (see

Supplemental Table 2). The mean ICC between the test and retest was 87.85%, indicating good to excellent reliability.

Assessment of the interrater reliability found that the mean test-retest ICC was 90.02% when participants were interviewed by the same rater and 86.58% when interviewed by a different rater, indicating good to excellent reliability for both approaches. Moreover, results of a two-sample t-test found no differences in the size of the discrepancy between test and retest scores for interviews conducted by the same ($M = 4.55$; $SD = 3.46$) or different ($M = 5.86$; $SD = 4.82$) raters ($t(126) = -1.67$, $p = .10$).

Phase 4: Assessing scale validity

Convergent validity

Participants and procedure

Convergent validity was assessed using the data from the 3,137 participants who were interviewed in the endline survey and the data from the 180 participants who were interviewed in the first round of assessment for the test-retest and reliability testing. The majority of the endline questions were developed based on prior instruments used in unpublished economic studies by the World Bank. Variables from the endline survey that were expected to be associated with the SSMHAS included: (1) life satisfaction as measured by a life satisfaction index (scored 0 to 100) that was rescaled from a composite of seven items rated on a 1–7 scale (i.e., “How satisfied are you with life as a whole?” “How satisfied are you with your education level?” etc.) and the present-day score on the ladder of life (Cantril, 1965); (2) optimism as measured by the ladder of life in five years and a one-item question asking whether the participant believes her daughter will have a better future than she does; (3) education, including whether the participant had attended school, started school by age seven, was currently in school, and could read or write a letter; (4) cognitive functioning as measured by the score of analytical ability on the Raven's Progressive Matrices (Raven & Court, 1993); (5) economic and employment status, including whether participants were engaged in any income generating activities, whether they had participated in a BRAC club, whether they think they will be working in two years, the quintile of their household wealth compared to that of the other participants in the endline survey, their confidence in

starting or running a business as measured by an entrepreneurship score (0 to 100), and daily stressors including the frequency (0 = never, 5 = everyday) of worry about getting a good job, finding a suitable husband, not having enough money, not having enough food, and being the victim of violence or theft; (6) trauma exposure, including the sum score of seven items assessing exposure to war violence pre 2006 and the sum score of seven items assessing exposure to the recent conflict since December 2013; and (7) social support as measured by the number of friends the participants has, whether the participant has someone in the community to stay with or borrow money or ask for help from, and whether the participant's mother, father, or both parents have died.

Variables from the reliability and validity sample that were expected to be associated with the SSMHAS included: (1) mental health as measured by the 12-item General Health Question (GHQ-12) scored as the sum of Likert scale items ranging from 0 (Never) to 3 (Always), where higher scores indicate more mental health problems (Piccinelli et al., 1993); and (2) education including whether the participant had attended school, started school by age seven, was currently in school, could read or write a letter, and the highest level of education they had completed from none to post-secondary.

Analyses

To assess the correlation between the continuous SSMHAS scores and variables that were hypothesized to be associated with the scale, linear regressions were run, with each hypothesized variable regressed on the mean score of the SSMHAS, with adjustment to the standard errors to account for the clustering in the sampling plan. Age was significantly associated with the mean score on the endline scale, and so it was included as a covariate in all regressions. Although half of the endline sample was randomized to the AGI intervention, the intervention was not significantly associated with the scale score, and so it was not included as a potentially confounding covariate.

Results

Scores on the GHQ-12 in the reliability sample ranged from 5 to 22, with a mean score of 14.47 (SD = 2.98), suggesting a potentially high burden of mental health symptoms since the highest suggest cut-off for the GHQ-12 is 3 (Goldberg et al., 1998). Complete results of linear regression being associated

with scores on the SSMHAS are presented in Table 3. Results indicated that scores on the SSMHAS were significantly associated with most of the hypothesized independent variables in the expected direction, including being negatively associated with cognitive functioning ($B = -0.068, p < .001$) and all of the measures of education ($p < .001$), and being positively associated with mental health problems as measured by the GHQ-12 ($B = 0.038, p < .001$), all of the daily stressor items ($ps < .001$) and recent conflict exposure ($B = 0.049, p < .001$).

Table 3
Linear Regressions Predicting Scores on the SSMHAS

Predictor	Endline Sample (N = 3137)*					
	n	%	B	95% CI	P	
Demographics						
Age, M(SD)	22.04	4.03	0.005	0.00	0.01	0.025
Have children	1805	57.56	-0.015	-0.05	0.02	0.448
In AGI intervention	1,561	49.76	-0.046	-0.14	0.05	0.320
Life satisfaction						
Where are you on the ladder of the life now, M (SD)	5.22	2.46	0.013	0.00	0.02	0.016
Life satisfaction index, M(SD)	55.49	18.19	0.002	.001	0.003	0.005
Optimism about the future						
Ladder of life in 5 years, M(SD)	6.99	2.40	0.004	-0.01	0.01	0.423
Will your daughter have a better future than yours?	2,779	88.59	0.056	-.03	.14	0.203
Education						
Ever attended school	2,503	79.79	-0.168	-0.23	-0.11	<.001
Started school on time	1,513	48.23	-0.109	-0.15	-0.06	<.001
Currently in school	820	26.14	-0.027	-0.07	0.02	0.254
Can read and write a letter	1,919	61.17	-0.082	-0.12	-0.04	<.001
Cognitive functioning						
Analytical ability (Raven's Matrices), M(SD)	3.18	1.38	-0.068	-0.08	-0.05	<.001
Economic status and employment						
Engaged in any income generating activity	2,109	67.23	0.079	0.03	0.13	0.001
Think you will be working in two years	2,488	79.34	0.211	0.16	0.26	<.001
Entrepreneurship score, M(SD)	63.43	21.69	0.001	0.00	0.002	0.006
Participated in a BRAC club	262	8.35	0.068	-0.02	0.16	0.125
Household wealth quintile, M (SD)	2.95	1.43	0.047	0.03	0.06	<.001
Daily stressors						
Frequency of worry about (0=Never, 5=Everyday):						
Not getting a good job, M(SD)	3.57	1.83	0.026	0.01	0.04	<.001
Not finding a suitable husband	3.18	1.97	0.050	0.04	0.06	<.001
Family not having enough money	3.22	1.72	0.040	0.03	0.05	<.001
family being the victim of violence or theft	2.09	1.57	0.078	0.06	0.10	<.001
Worried about your family not having enough food (n, %)	2,016	64.33	0.080	0.03	0.13	.001
Trauma exposure						
Pre 2006 war exposure, M(SD)	1.77	1.84	0.007	-0.01	0.02	0.364
Since Dec 2013 conflict exposure	1.38	1.91	0.049	0.04	0.06	<.001
Social support						
Number of friends, M(SD)	3.55	2.40	0.005	-0.003	0.01	0.193
Does not have a mother, father, or both	404	12.88	-0.044	-0.12	0.03	0.265

Have someone to help you	2,402	76.59	0.131	0.08	0.19	<.001
Validity and Reliability Sample (N=180)						
Predictor	n	%	B	95% CI		P
Demographics						
Age, M(SD)	25.22	6.72	-0.001	-0.007	0.005	0.630
Have children						
Mental health symptoms						
General Health Questionnaire, M(SD)	14.47	3.93	.038	.031	.045	<.001
Education						
Ever attended school	163	90.56	-0.052	-0.30	0.20	0.669
Started school on time	129	71.67	0.037	-0.07	0.15	0.501
Currently in school	40	22.22	-0.106	-0.26	0.05	0.169
Highest level of education completed, M (SD)	7	3.78	-0.011	-0.03	0.008	0.245
Can read and write a letter	124	68.89	-0.083	-0.22	0.06	0.238

Note. *adjusted for age

However, scores on the SSMHAS were also significantly associated with some indicators in the opposite direction of what was expected (see Table 2). Specifically, life satisfaction (including current status on the ladder of life ($B = 0.32, p = .016$) and the life satisfaction index ($B = 0.05, p = .005$)), all of the measures of economic status and employment including engaging in an income generating activity ($B = 1.90, p = .001$), thinking you will be working in two years ($B = 5.06, p < .001$), entrepreneurship ($B = 0.03, p = .006$), and household wealth quintile ($B = 1.14, p < .001$), and having someone in the community who can help you ($B = 3.13, p < .001$) were significantly associated in the opposite direction such that greater life satisfaction, economic status and employment, and having someone to help associated with higher scores on the SSMHAS and therefore more mental health problems. Due to the relatively low symptom severity in this population, these analyses were rerun on participants with the top quarter and bottom quarter of the AGI scale to ensure that association patterns remained the same even among those who were the most and least distressed. Results indicated that the direction of these associations remained the same.

Concurrent validity

To assess whether scores on the scale correlate with self and other reports of having a mental health problem, 22 of the leaders involved in administering the AGI program, and who knew the participants, were asked to indicate whether participants from the reliability and validity sample had each of the identified mental health problems (e.g., would you consider [name of participant] to be traumatized (kafu)?” For each mental health problem, leaders could rate a participant as definitely having the problem, somewhat having the problem, definitely not having the problem, or don’t know. For this study, to be rated as having a problem based on leader report, participants had to be rated by a leader as “definitely having the problem.” The total number of problems that each participant had based on leader report was calculated by totaling the number of problems that leaders reported as “definitely yes.” Therefore, the total number of problems could range from 0 to 6. Don’t know responses were rated as missing. Not all leaders were located, and so leader reports are only available for 136 of the 180 participants in the reliability and validity sample.

Six of the items on the SSMHAS used the name of the problem as the prompt (e.g., “How often do you feel shame (fadia)?”), and therefore participants self-rated how often they experienced each of the six mental health conditions 1 (Never), 2 (Rarely), 3 (Frequently), and 4 (Always). Participants who rated themselves as having a problem “Frequently” or “Always” were rescored into a binary variable as having that problem based on self-report. The total number of problems that each participant had based on self-report was calculated by totaling the number of problems that participants self-reported that they had “Frequently” or “Always” on the SSMHAS. Therefore, the total number of self-reported problems each participant had could range from 0 to 6.

Analyses

To assess the concordance between participant self-report and leader report for each of the mental health problems, McNemar's chi-square tests were run. In addition, to assess the association between the number of problems reported by participants and by leaders, a paired sample t-test was run. Finally, to assess the association between scores on the SSMHAS, participant self-report of mental health problems, and leader report of participant mental health problems, Spearman correlations were calculated.

Results

The 22 AGI group leaders rated 76 (56%) of the 136 participants they were asked about as definitely having at least one of the six problems (see Table 4). Of the 180 participants in the reliability and validity sample, 80 (44.44%) self-rated as having a problem “frequently” or “always.” Concurrent validity of the SSMHAS was assessed by determining whether leader reports of mental health problems were significantly associated with participant self-rated scores on the scale (see Tables 3 and 4). Results found that leader reports of participants being stressed out ($r_s = .30, p = .002$), having any of the six problems ($r_s = .19, p = .04$), and the total number of problems reported for each participant ($r_s = .21, p = .03$) were all positively associated with participant scores on the assessment scale.

Table 4*Leader and Self-Reports of Having Mental Health Problems*

	Leader's report (N=136)			Participant self-report (N=180)			Comparison (N=136)		
	N	n	%	N	n	%	N	Chi-square	p
Pain in heart (<i>waja gelba</i>)	125	18	14.4	180	40	22.22	104	0.73	0.729
Stressed up (<i>nefsiat</i>)	128	26	20.31	180	27	15	108	1.67	0.196
Shame (<i>fadia</i>)	134	30	22.39	180	27	15	114	2.54	0.111
Rudeness (<i>takian</i>)	136	32	23.53	180	17	9.44	115	5.16	0.023
Isolated or abandoned (<i>abau / hasud</i>)	135	7	5.19	180	4	2.22	114	0.20	0.653
Traumatized (<i>kafu</i>)	136	11	8.09	180	11	6.11	115	3.71	0.054
Yes to any problem	136	76	55.88	180	80	44.44	115	0.16	0.692
Number of problems, M (SD)	136	0.91	1.01	180	0.55	0.86	136	t(135)=3.28	0.001

Table 5*Spearman Correlations Between the SSMHAS and Leader and Participant Reports of Mental Health Problems*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Scale	---															
Leader Report																
2. Pain in heart	0.14	---														
3. Stressed up	0.30**	0.35***	---													
4. Shame	0.07	-0.05	0.09	---												
5. Rude	-0.01	-0.13	-0.04	-0.04	---											
6. Isolated	0.00	0.10	0.16	0.12	0.27**	---										
7. Traumatized	0.06	-0.04	0.12	0.30***	0.09	0.19*	---									
8. Any problem	0.19*	0.36***	0.44***	0.48***	0.49***	0.21*	0.26**	---								
9. # of problems	0.21*	0.44***	0.57***	0.53***	0.46***	0.39***	0.42***	0.92***	---							
Participant report																
10. Pain in heart	0.28***	-0.08	-0.11	0.12	-0.14	-0.13	0.07	-0.13	-0.11	---						
11. Stressed up	0.46***	-0.02	0.12	-0.12	-0.07	-0.11	0.05	0.00	-0.03	0.11	---					
12. Shame	0.25***	0.07	-0.05	0.15	0.02	0.12	-0.03	0.07	0.07	0.19	0.00	---				
13. Rude	0.20**	0.05	0.03	-0.09	0.21*	0.06	-0.09	0.12	0.09	0.10	0.02	0.13	---			
14. Isolated	0.25***	0.13	0.17	0.04	-0.09	-0.04	0.34***	0.03	0.12	0.01	0.25***	0.04	-0.05	---		
15. Traumatized	0.28***	0.11	0.06	-0.06	-0.06	-0.07	0.18	0.07	0.05	-0.02	0.22**	0.09	0.00	0.12	---	
16. Any problem	0.56***	0.00	0.04	0.03	-0.03	-0.07	0.11	0.04	0.03	0.60***	0.47***	0.47***	0.36***	0.17*	0.29***	---
17. # of problems	0.63***	0.07	0.09	0.05	-0.02	-0.03	0.13	0.07	0.08	0.65***	0.56***	0.51***	0.37***	0.26**	0.38***	0.96***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Results of a paired t-test indicated that leaders rated participants as definitely having more problems ($M = 0.91$, $SD = 1.01$) than participants rated themselves as having frequently or always ($M = 0.55$, $SD = 0.86$; $t(135) = 3.28$, $p < .001$). Results of McNemar's chi-square tests between leader and participant endorsement of participant mental health problems found that ratings of Rudeness (takian) (McNemar's $\chi^2 = 13.37$, $p < .001$), Shame (fadia) (McNemar's $\chi^2 = 4.80$, $p = .029$), and endorsement of any of the problems (McNemar's $\chi^2 = 6.33$, $p = .012$) differed between leader ratings and participant self-ratings. However, there were no differences by rater type for the other problems, suggesting agreement; see Table 4 for complete results. In addition, there was a significant correlation between participants' reports of having rudeness and leader reports of participants having rudeness ($r_s = .21$, $p = .02$). In addition, participant report of isolation was significantly associated with leader reports of being traumatized ($r_s = .34$, $p < .001$). See Table 5 for correlation matrix.

To further assess the construct validity of the leader reports, analyses assessed whether leader scores were associated with mental health symptoms measured by the GHQ-12. Results indicated that leader reports of participants having pain in the heart ($r_s = .21$, $p = .02$), feeling stressed out ($r_s = .21$, $p = .02$), having any problem ($r_s = .22$, $p = .01$), and the total number of problems ($r_s = .32$, $p < .001$) were all significantly associated with higher scores on the GHQ-12.

Discussion

The goal of this study was to develop and validate a measure that would capture variation in local idioms of distress that may be commonly experienced by women in South Sudan. Six idioms of distress were identified and defined and the SSMHAS was developed to assess these symptoms. Although the SSMHAS could have assessed multiple domains of mental health, the scale had strong internal reliability with one factor, suggesting that the assessment of all six individual idioms of distress collectively assessed one underlying construct of distress. Notably, previous research has identified "overthinking" as an idiom of distress in South Sudan, which was identified in this study as one symptom of "Stressed out" (Adaku et al., 2016). In addition, despite having low variability in responses, the scale had strong test-retest and interrater reliability, and the overall scores were very stable, with participants having nearly

identical scores at test and retest, suggesting that the construct the scale was stable over an average of 12 days later.

Despite the somewhat truncated range of the SSMHAS, results assessing convergent validity still found that participants who reported more symptoms on the SSMHAS had worse mental health as assessed by the GHQ-12, worse cognitive functioning, more daily stressors, and more conflict exposure than those who reported fewer mental health symptoms, indicating good convergent validity for the scale. Additionally, although some associations were in the opposite direction than were originally hypothesized, they may still indicate strong convergent validity. For example, although participants who reported currently engaging in more economic activities, planning to engage in economic activities in two years, and having higher entrepreneurship also reported more mental health symptoms, these factors were also associated with more worries about money, not having a job, not having clothes to wear, not finding a husband, and experiencing violence. In addition, they were negatively associated with having education. Therefore, engagement in economic strengthening activities appears to be an indication of more concern about meeting daily needs, and its positive association with mental health problems is consistent. However, the results indicating that higher current life satisfaction and having someone to help you is associated with higher scores on the AGI scale are not consistent. It may be that since few of the participants were highly symptomatic, the lack of range in the data is not adequately identifying the association between life satisfaction or having someone to help you and symptoms. This may be particularly true since there was greater range in life satisfaction and having someone help than there was in scores on the AGI scale.

Results assessing whether observer (leader) reports of mental health problems were associated with scores on the SSMHAS also indicated concurrent validity. Leader reports of rudeness, trauma, and of participants having some mental health problems were associated with participant ratings on the scale. These results are consistent with the idea that observers may be able to identify externalizing symptoms or general distress, but may not be able to identify specific internalizing problems. The results suggest

that the scale was measuring symptoms that may be identifiable to others and may be negatively impacting functioning.

The SSMHAS scale was designed to measure local idioms of distress and culturally appropriate indicators of general well-being and was not intended to assess psychopathology or clinical levels of mental health problems. Nonetheless, scores on the SSMHAS were positively and linearly associated with scores on the GHQ-12, which was designed to screen for psychiatric disorders and has been validated in several countries including, but not limited to, Nigeria (Gureje & Obikoya, 1990), Kenya (Abubakar & Fischer, 2012), Spain (Sánchez-López Mdel & Dresch, 2008), Iran (Montazeri et al., 2003), Poland (Makowska et al., 2002), and Brazil (De Jesus Mari & Williams, 1985). Scores on the GHQ-12 ranged from 5 to 22, with an average score of 14 (SD = 3), which indicates elevated score across validation samples from other settings (Goldberg et al., 1998). This elevation is in line with other studies from South Sudan that found elevated rates of depression and PTSD using non-locally validated measures (Ayazi et al., 2012; Karunakara et al., 2004; Ng et al., 2017; Roberts et al., 2009). Additionally, scores on the SSMHAS were associated with several indicators of worry about basic needs and safety and with trauma exposure which have been found to be strongly associated with mental health in similar conflict-exposure and resource-constrained settings (Miller and Rasmussen, 2010; Miller et al., 2008). Therefore, we hypothesize that scores on the SSMHAS would be correlated with psychiatric distress in this population.

The apparent discrepancy in reported symptoms on the SSMHAS and the GHQ-12 highlights the importance of developing and utilizing locally derived measures rather than relying solely on scales developed in other populations and contexts. Measures that are not locally developed and validated have been found to lead to substantial under- or over-reporting of mental health symptoms (Madigoe et al., 2017). With regards to the SSMHAS and the GHQ-12, either measure could be accurately, over-, or under-identifying mental health problems. However, while the SSMHAS and the GHQ-12 were hypothesized to correlate, they were also designed to assess different constructs. The SSMHAS was specifically developed to assess locally derived and locally meaningful symptoms and may therefore have more validity for this population in this context than the GHQ (Miller et al., 2009; Ng et al., 2014).

However, the GHQ may be identifying symptoms that would otherwise not be conceptualized as problematic and/or related to mental health in South Sudan, but that may still result in functional impairment and distress. Unfortunately, there is limited literature on the development and validation of mental health scales outside of Western populations (Kohrt et al., 2014) and few studies that assess mental health prevalence utilizing locally validated scales (Charlson et al., 2019; Ng et al., 2020). The DIME approach provides a pragmatic and feasible way for non-mental health specialists to develop and validate culturally and contextually appropriate mental health tools (Applied Mental Health Research Group, 2013) and increased use of this and other approaches would strengthen understanding of mental health globally.

Overall, this scale appears to perform well psychometrically and adds to the literature by including common local idioms of distress that are relevant to, and identifiable in, South Sudan. Therefore, this scale may be used more easily and readily in the local context than scales developed in other settings that have been translated into the local language. In addition, the development of the scale may provide an example or roadmap for development agencies, NGOs, or other organizations working in environments where mental health measures have not yet been developed and validated to create and validate measures that are relevant to their populations using the teams that they have on the ground. In this way, the role of mental health in development settings can be more rapidly assessed and understood, even in places experiencing active conflict.

This study has several limitations. First, participants were drawn from those who participated in the AGI RCT, and the Free Listing, Key Informant interviews, and reliability sample were only collected in Juba. The data are therefore not representative of the population and cannot be generalized to the larger South Sudanese population. The scale requires further assessment on a more representative sample. In addition, the sample was fairly high-functioning and non-clinical as indicated by the low scores on the SSMHAS. While these results are not surprising given that the sample was selected from participants who were willing and able to participate in an economic strengthening intervention, it does decrease the generalizability of the scale since it was not validated on participants with the full range of common

mental health problems. Given this, the scale is appropriate to use for the assessment of mental health in the context of an economic development, educational, or physical health intervention, but is not appropriate for use clinically without further validation. Additionally, although the DIME manual states that for test-retest evaluation, re-administration of the measure should occur one to three days after the first administration (Center for Refugee & Disaster Response, 2016), this was infeasible for this study and so participants were reassessed from 0 to 36 days after the first administration, with an average interval of 12 days. Finally, given the ongoing conflict, we were unfortunately unable to have the local partners participate in the write-up of this study.

Despite these limitations, this study is the first study to develop and validate a measure of mental health symptoms that reflect the local idioms of distress in South Sudan, a country that has been decimated by war and may have a high burden of mental health concerns. Moreover, this article describes a methodology of measure development that may be used by other NGOs operating in fragile or conflict-afflicted settings who want to measure mental health in their communities but may not have been able to conduct a large-scale validity study due to security or resource constraints. This pragmatic methodological approach could be used to close the gap in our knowledge about local idioms of distress and may provide the broader mental health field with sorely needed locally valid mental health scales for some of the most vulnerable populations around the world.

References

- Abubakar, A., Fischer, R. (2012). The factor structure of the 12-item general health questionnaire in a literate Kenyan population. *Stress and Health, 28*(3), 248-254. <https://doi.org/10.1002/smi.1420>
- Adaku, A., Okello, J., Lowry, B., Kane, J. C., Alderman, S., Musisi, S., Tol, W. A. (2016). Mental health and psychosocial support for South Sudanese refugees in northern Uganda: A needs and resource assessment. *Conflict and Health, 10*(1), 18. <https://doi.org/10.1186/s13031-016-0085-6>
- Applied Mental Health Research Group. (2013). Design, implementation, monitoring, and evaluation of mental health and psychosocial assistance programs for trauma survivors in low resource countries: A user's manual for researchers and program implementers. http://hopkinshumanitarianhealth.org/assets/documents/VOT_DIME_MODULE1_FINAL.PDF
- Ayazi, T., Lien, L., Eide, A. H., Jenkins, R., Albino, R. A., Hauff, E. (2013). Disability associated with exposure to traumatic events: Results from a cross-sectional community survey in South Sudan. *BMC Public Health, 13*, 469. <http://www.biomedcentral.com/1471-2458/13/469>
- Ayazi, T., Lien, L., Eide, A. H., Ruom, M. M., Hauff, E. (2012). What are the risk factors for the comorbidity of posttraumatic stress disorder and depression in a war-affected population? A cross-sectional community study in South Sudan. *BMC Psychiatry, 12*, 175. <http://www.biomedcentral.com/1471-244X/12/175>
- Badri, A., Crutzen, R., Van den Borne, H. W. (2012). Exposures to war-related traumatic events and post-traumatic stress disorder symptoms among displaced Darfuri female university students: An exploratory study. *BMC Public Health, 12*, 603. <http://www.biomedcentral.com/1471-244X/12/175>
- Ballon, P., Duclos, J.-Y. (2015). Multidimensional poverty in Sudan and South Sudan. <https://ssrn.com/abstract=2592627>
- Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., Sulaiman, M. (2015). Women's empowerment in action: Evidence from a randomized control trial in Africa. *American Economic Journal: Applied Economics, 12*(1), 210–59
- Cantril, H. (1965). The pattern of human concerns. Rutgers University Press.
- Center for Refugee and Disaster Response. (2016). The DIME program research model: Design, Implementation, Monitoring, and Evaluation. http://www.jhsph.edu/research/centers-and-institutes/center-for-refugee-and-disaster-response/response_service/AMHR/dime/index.html
- Charlson, F., van Ommeren, M., Flaxman, A., Cornett, J., Whiteford, H., Saxena, S. (2019). New WHO prevalence estimates of mental disorders in conflict settings: A systematic review and meta-analysis. *Lancet (London, England), 394*(10194), 240-248. [https://doi.org/10.1016/s0140-6736\(19\)30934-1](https://doi.org/10.1016/s0140-6736(19)30934-1)
- De Jesus Mari, J., Williams, P. (1985). A comparison of the validity of two psychiatric screening questionnaires (GHQ-12 and SRQ-20) in Brazil, using relative operating characteristic (ROC) analysis. *Psychological Medicine, 15*(3), 651-659. <https://doi.org/10.1017/S0033291700031500>
- Friedrich, M. J. (2016). Psychological scars from conflict in South Sudan documented. *JAMA, 316*(11), 1141. <https://doi.org/10.1001/jama.2016.12651>

- Goldberg, D. P., Oldehinkel, T., Ormel, J. (1998). Why GHQ threshold varies from one place to another. *Psychological Medicine*, 28(4), 915-921. <https://doi.org/10.1017/s0033291798006874>
- Gureje, O., Obikoya, B. (1990). The GHQ-12 as a screening tool in a primary care setting. *Social Psychiatry and Psychiatric Epidemiology*, 25(5), 276-280. <https://doi.org/10.1007/BF00788650>
- Hinton, D. E., Lewis-Fernández, R. (2010). Idioms of distress Among trauma survivors: Subtypes and clinical utility. *Culture, Medicine, and Psychiatry*, 34(2), 209-218. <https://doi.org/10.1007/s11013-010-9175-x>
- Karunakara, U. K., Neuner, F., Schauer, M., Singh, K., Hill, K., Elbert, T., Burnham, G. (2004). Traumatic events and symptoms of post-traumatic stress disorder amongst Sudanese nationals, refugees and Ugandans in the west Nile. *African Health Sciences*, 4(2), 83-93. PMID: 15477186; PMCID: PMC2141616.
- Kohrt, B. A., Rasmussen, A., Kaiser, B. N., Haroz, E. E., Maharjan, S. M., Mutamba, B. B., de Jong, J. T., Hinton, D. E. (2014). Cultural concepts of distress and psychiatric disorders: Literature review and research recommendations for global mental health epidemiology. *International Journal of Epidemiology*, 43(2), 365-406. <https://doi.org/10.1093/ije/dyt227>
- Lien, L., Hauff, E., Martinez, P., Eide, A. H., Swartz, L., Ayazi, T. (2016). Alcohol use in South Sudan in relation to social factors, mental distress and traumatic events. *BMC Public Health*, 16(1), 937. <https://doi.org/10.1186/s12889-016-3605-9>
- Madigoe, T., Burns, J., Zhang, M., Subramaney, U. (2017). Towards a culturally appropriate trauma assessment in a South African Zulu community. *Psychological Trauma*, 9(3), 274-281. <https://doi.org/10.1037/tra0000231>
- Makowska, Z., Merez, D., Mościcka, A., Kolasa, W. (2002). The validity of general health questionnaires, GHQ-12 and GHQ-28, in mental health studies of working people. *International Journal of Occupational Medicine and Environmental Health*, 15(4), 353-362. PMID: 12608623.
- Meffert, S. M., Musalo, K., Abdo, A. O., Alla, O. A., Elmakki, Y. O., Omer, A. A., ... Marmar, C. R. (2010). Feelings of betrayal by the united nations high commissioner for refugees and emotionally distressed Sudanese refugees in Cairo. *Medicine, Conflict, and Survival*, 26(2), 160-172. <https://doi.org/10.1080/13623699.2010.491395>
- Miller, K., Omidian, P., Rasmussen, A., Yaqubi, A., Daudzai, H. (2008). Daily stressors, war experiences, and mental health in Afghanistan. *Transcultural Psychiatry*, 45(4), 611-638. <https://doi.org/10.1177/1363461508100785>
- Miller, K. E., Omidian, P., Kulkarni, M., Yaqubi, A., Daudzai, H., Rasmussen, A. (2009). The validity and clinical utility of post-traumatic stress disorder in Afghanistan. *Transcultural Psychiatry*, 46(2), 219-237. <https://doi.org/10.1177/1363461509105813>
- Miller, K. E., Rasmussen, A. (2010). War exposure, daily stressors, and mental health in conflict and post-conflict settings: Bridging the divide between trauma-focused and psychosocial frameworks. *Social Science & Medicine* (1982), 70(1), 7-16. <https://doi.org/10.1016/j.socscimed.2009.09.029>
- Ministry of Health and National Bureau of Statistics. (2013). South Sudan Household Survey 2010. Final report.

<https://reliefweb.int/sites/reliefweb.int/files/resources/SHHS%20II%20Report%20Final.pdf>
(Accessed 16 Nov 2021).

- Montazeri, A., Harirchi, A. M., Shariati, M., Garmaroudi, G., Ebadi, M., Fateh, A. (2003). The 12-item general health questionnaire (GHQ-12): Translation and validation study of the Iranian version. *Health and Quality of Life Outcomes*, 1(1), 66. <https://doi.org/10.1186/1477-7525-1-66>
- Ng, L. C., Kanyanganzi, F., Munyanah, M., Mushashi, C., Betancourt, T. S. (2014). Developing and validating the youth conduct problems scale-Rwanda: A mixed methods approach. *PLoS One*, 9(6), e100549. <https://doi.org/10.1371/journal.pone.0100549>
- Ng, L. C., Lopez, B., Pritchard, M., Deng, D. (2017). Posttraumatic stress disorder, trauma, and reconciliation in South Sudan. *Social Psychiatry and Psychiatric Epidemiology*, 52(6), 705-714. <https://doi.org/10.1007/s00127-017-1376-y>
- Ng, L. C., Stevenson, A., Kalapurakkel, S. S., Hanlon, C., Seedat, S., Harerimana, B., Koenen, K. C. (2020). National and regional prevalence of posttraumatic stress disorder in sub-Saharan Africa: A systematic review and meta-analysis. *PLoS Medicine*, 17(5), e1003090. <https://doi.org/10.1371/journal.pmed.1003090>
- Piccinelli, M., Bisoffi, G., Bon, M. G., Cunico, L., Tansella, M. (1993). Validity and test-retest reliability of the Italian version of the 12-item general health questionnaire in general practice: A comparison between three scoring methods. *Comprehensive Psychiatry*, 34(3), 198-205. [https://doi.org/10.1016/0010-440x\(93\)90048-9](https://doi.org/10.1016/0010-440x(93)90048-9)
- Raven, J. C., Court, J. H. (1993). Manual for raven progressive matrices and vocabulary scales. Psychological Corporation.
- Roberts, B., Damundu, E. Y., Lomoro, O., Sondorp, E. (2009). Post-conflict mental health needs: A cross-sectional survey of trauma, depression and associated factors in Juba, southern Sudan. *BMC Psychiatry*, 9, 7. <https://doi.org/10.1186/1471-244x-9-7>
- Sánchez-López Mdel, P., Dresch, V. (2008). The 12-item general health questionnaire (GHQ-12): Reliability, external validity and factor structure in the Spanish population. *Psicothema*, 20(4), 839-843. PMID: 18940092.
- Schweitzer, R., Melville, F., Steel, Z., Lacherez, P. (2006). Trauma, post-migration living difficulties, and social support as predictors of psychological adjustment in resettled Sudanese refugees. *Australian and New Zealand Journal of Psychiatry*, 40(2), 179-187. <https://doi.org/10.1111/j.1440-1614.2006.01766.x>
- Tol, W. A., Augustinavicius, J., Carswell, K., Brown, F. L., Adaku, A., Leku, M. R., García-Moreno, C., Ventevogel, P., White, R. G., van Ommeren, M. (2018). Translation, adaptation, and pilot of a guided self-help intervention to reduce psychological distress in South Sudanese refugees in Uganda. *Global Mental Health*, 5, e25. <https://doi.org/10.1017/gmh.2018.14>
- United Nations Development Programme. (2018). Human development report 2018. http://hdr.undp.org/sites/default/files/2018_human_development_statistical_update.pdf
- Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the

global burden of disease study 2010. *Lancet (London, England)*, 382(9904), 1575-1586.
[https://doi.org/10.1016/s0140-6736\(13\)61611-6](https://doi.org/10.1016/s0140-6736(13)61611-6)

World Bank. (2016). Adolescent Girls Initiative. <http://www.worldbank.org/en/programs/adolescent-girls-initiative>

Supplemental Table 1. SSMHAS Item Level Factor Loadings

Item	Loading on Factor 1	Uniqueness
Have misunderstandings	0.40	0.84
Have a quick temper	0.53	0.72
Get into fights or quarrels	0.55	0.69
Feel rude (<i>takian</i>)	0.55	0.70
Have many thoughts	0.49	0.76
Have sleepless nights	0.51	0.74
Feel unsettled	0.60	0.63
Feel stressed up (<i>nefsiat</i>)	0.54	0.70
Feel lonely	0.55	0.70
Feel unwanted	0.60	0.64
Feel useless	0.59	0.66
Feel isolated or abandoned (<i>hasud or abau</i>)	0.58	0.66
Feel weak	0.48	0.77
Feel sorrowful	0.53	0.72
Feel like crying	0.60	0.64
Feel pain in heart (<i>waja gelba</i>)	0.56	0.69
Are you easily startled	0.61	0.63
Think about bad things that happened in the past	0.53	0.72
Do bad things in the past make you act crazy	0.62	0.62
Feel traumatized (<i>kafu</i>)	0.61	0.62
Fear that others are talking about you	0.56	0.69
Feel shy	0.52	0.73
Hide from others	0.50	0.75
Feel shame (<i>fadia</i>)	0.43	0.82

Supplemental Table 2. Test-Retest and Interrater Reliability Intraclass Correlation (ICC) Scores

<i>Item.</i> How often do you...	ICC agreement		
	Full sample	Same rater	Different raters
get into fights or quarrels?	89.15%	84.31%	92.21%
have misunderstandings?	90.62%	86.00%	93.51%
have a quick temper?	76.74%	82.35%	74.03%
feel rude (takian)?	86.82%	90.20%	84.42%
have many thoughts?	72.87%	80.39%	68.83%
how often do you have sleepless nights?	79.07%	86.27%	74.03%
feel unsettled?	95.35%	96.08%	94.81%
feel stressed up?	79.07%	80.39%	79.22%
feel lonely?	93.80%	96.08%	93.51%
feel unwanted?	96.12%	98.04%	94.81%
feel useless?	96.12%	96.08%	96.10%
feel isolated or abandoned (hasud or abau)	100.00%	100.00%	100.00%
feel weak?	86.82%	90.20%	84.42%
feel sorrowful?	93.02%	90.20%	94.81%
feel like crying?	81.40%	78.43%	84.42%
feel pain in heart (waja gelba)?	72.09%	76.47%	68.83%
are you easily startled?	89.15%	88.24%	89.61%
think about bad things that happened in the past?	84.50%	94.12%	77.92%
do bad things in the past make you act crazy?	98.45%	98.04%	98.70%
feel traumatized (kafu)?	88.37%	92.16%	87.01%
fear that others are talking about you?	93.02%	94.12%	92.21%
feel shy?	93.02%	96.08%	90.91%
hide from others?	97.67%	98.04%	97.40%
feel shame (fadia)?	75.19%	88.24%	66.23%
Mean	87.85%	90.02%	86.58%