Assessment of Psychosocial Problems and Impaired Functioning among Conflict Affected Adults in Aceh, Indonesia: Testing of a Locally-Adapted Psychosocial Assessment Instrument and Screening into an Intervention Program

Conducted by
International Catholic Migration Committee/Indonesia (ICMC)
Rehabilitation Action for Torture Victims in Aceh (RATA)
The Applied Mental Health Research Group
Johns Hopkins Bloomberg School of Public Health

Authors
Judy Bass
Melissa Azur
Bhava Poudhal
Paul Bolton

January, 2008

TABLE OF CONTENTS

EXECUTIVE SUMMARY

ACKNOWLEDGEMENTS

INTRODUCTION

This report describes the testing and initial utilization of an instrument to assess psychosocial problems among adults affected by the 30-year conflict between the Free Aceh Movement (GAM) and the Indonesian government forces in Aceh For the current study, testing of the psychosocial assessment instrument consisted of assessing the instrument's local validity and reliability among the target population. Once we were assured that the instrument was valid, it was used to screen people into the subsequent intervention evaluation study, which is currently in progress.

This report briefly describes the background to the activities described here, placing them in the wider context of ongoing technical collaboration between the International Catholic Migration Commission Indonesia (ICMC) and Johns Hopkins University (JHU). The report describes the methods used to test the instrument and the results of these tests and its initial use as a screening tool for the intervention study. Conclusions based on the results are also included, as well as specific recommendations for future activities. Finally, the report ends with a series of appendices containing the final instrument and associated interview materials. There is also an appendix with explanation of some of the technical terms used in the report referring to important elements of reliability and validity.

BACKGROUND

Situation in Aceh
In the context of Aceh, violence has occurred for over 30 years, during the conflict between The Free Aceh Movement (GAM) and the Indonesian Government. As GAM tried to struggle for independence, and the Indonesian government tried to curb it through military operations since the mid 1970s, people of Aceh have experienced and witnessed violence. During the Suharto regime, Aceh faced a lot of violence, and was hopeful after the fall of Suharto in 1998, but the intimidation, arrest, torture and arbitrary killings continued. 2004 saw another militarization of Aceh, which continued until the devastating tsunami in Dec 2005. After the tsunami, both GAM and the Government of Indonesia entered into a truce, which resulted in a peace settlement in September 2006.

The site targeted for this project, Bereuen, was one of the most affected districts. It was considered as one of the strongholds of GAM and was frequently attacked by the military. GAM had their hide outs in the forests, however villages around the forest areas were prime targets for the military. All of the villages participating in this evaluation study were highly affected by the conflict. Entire villages experienced torture first hand, or were family members or witnesses of torture and arbitrary killings.

These activities described here (testing and initial utilization) form part of a planned series of field-based activities to inform the design, monitoring, and evaluation (DME) of the ICMC counseling program for conflict affected adults. These DME activities are being conducted by ICMC in collaboration with faculty from Johns Hopkins University. The collaboration is supported by the Victims of Torture Fund (VTF) at USAID.

The DME process consists of the following stages:

1. Qualitative study of local concepts of function, and of the problems affecting the target population (including identification of locally important psychosocial problems).
2. Development of a locally appropriate quantitative instrument (questionnaire) to assess the major psychosocial problems and local indicators of functioning emerging from the qualitative study.
3. Evaluation of the acceptability and clarity of the instrument through a pilot study among the target population, with subsequent revision of the instrument based on the results.
4. Testing of the reliability and validity of the final instrument.
5. Use of the final version of the instrument to conduct baseline assessments among adults recruited to the program.
6. Repeat use of the instrument after participation in the program, to assess program impact.

This report describes the process and results of stages 4 and 5, described above. The report on the previous qualitative study which constitute stage 1 (and upon which the current work has been based) is currently being drafted by ICMC. The report on the development and piloting of the instrument (stages 2 & 3) is currently being revised and a draft is available from this report's authors. Details on the background to ICMC's work in Indonesia, the overall DME project and methodology; its rationale, theoretical basis, and explanations of the technical concepts, are described in detail in these earlier reports, therefore much of this information has been omitted from this report. The contents here are limited to a description of the methods used to testing the instrument's reliability and validity in the field and its utilization as a baseline screener.

OBJECTIVES

The objectives of the work described here are:
1. To test the validity and reliability of this instrument among adults in 6 conflict affected villages in Aceh, Indonesia.
2. To screen adults in each village using the validated version of the instrument, to identify those with significant psychosocial problems that might benefit from ICMC's counseling program.
3. To continue to build the capacity of ICMC in this type of applied research.

METHODS

Study Sites

This study took place in 6 villages around the town of Bireuen in Aceh, Indonesia in August 2007. These villages were selected from the 15 target villages identified in RATA's initial program proposal to ICMC. All 15 villages were highly affected by the conflict, with entire villages being exposed to torture either directly, through family members or through witnessing of torture and arbitrary killings. Prior to this phase of the process, 4 of those 15 villages already received ICMC's counseling program as part of the RATA training and piloting program. From the remaining 11 villages, RATA was asked to choose 6 to be involved in the evaluation study. To select the 6 villages, RATA was asked to identify pairs of villages that were similar in terms

of population size and distance from the town, with the intent that one of each pair would receive the intervention and the other would be the control. In the end, it was not possible to match on population size, so the focus was on pairing villages that were in a similar area (i.e. both North of the main town) and similar distance from the main town of Bireuen. Table 1 presents the basic population structures of the 6 villages. In terms of distance from the town of Bireuen, age and gender structure, the pairings appear adequate. In terms of population size, there was some variation between the village pairings: Blang Ketumba and Abeuk Usong both are approximately 50% larger in population size compared with their paired villages, Blang Gandai and Batee Raya. All 6 of these villages received the same socialization program as an introduction to the program (an example of the socialization process is presented in Appendix A). Determination of which villages would receive the intervention and which would be controls was not done until all of the screening interviews were completed in all the villages, so as not to bias the interviewing and recruitment procedures.

TABLE 1: Demographics of the Pairs of Study Villages*

Village	Blang Gandai	Blang Ketumba		Batee Raya	Abeuk Usong		Abeuk Budi	Peruaden
Intervention Status**	I	C		I	C		I	C
Population Size	1051	1647		697	1184		682	513
Males (%)	536 (51)	804 (49)		316 (45)	565 (48)		323 (47)	242 (47)
Females (%)	515 (49)	843 (51)		381 (55)	619 (52)		359 (53)	271 (53)
Age Breakdown								
0-14 years (%)	351 (33)	519 (32)		260 (37)	387 (33)		233 (34)	164 (32)
15-64 years (%)	660 (63)	1074 (65)		389 (56)	737 (62)		415 (61)	315 (61)
65+ years (%)	40 (4)	54 (3)		48 (7)	60 (5)		34 (5)	34 (7)

* Population data provided by community leaders.
** I stands for Intervention village; C stands for Control village

Once all of the baseline interviews were completed, a discussion was had with the RATA counselors to choose which villages should receive the intervention and which would be controls. The counselors were not informed about any results from the data analysis prior to these conversations so as not to bias their choices. Two of the villages (Abeuk Budi and Blang Gandai) were chosen to receive the intervention now as they had more difficult roads that would be harder to navigate in the January heavy rainy season, compared with their matched villages of Blang Ketumba and Peruaden. In addition, the RATA counselors indicated that many of the people in Peruaden would spend these first months in the forests for work but would more likely stay in the villages starting in December, again because of the heavy rains. For the two larger villages, Batee Raya and Abeuk Usong, the RATA counselors indicated that it did not matter which received the intervention first and so they chose Batee Raya. The intervention villages were to receive the counseling first. The control villages would not receive the intervention until January, once the intervention had been completed in the intervention villages and the follow-up assessment was done.

<u>Scale Development</u>

For the assessment of psychosocial symptoms, we used an adapted version of the Hopkins Symptom Checklist (HSCL-25) which includes two subscales, one for assessing depression and the other for assessing anxiety. We also adapted a WHO somatic scale, as well as 14 symptom items specific to this Acehnese population, based on the qualitative data and not already represented in the existing scales.

To determine the best way to represent the symptoms in scale format, an exploratory factor analysis was then conducted with all 44 symptoms. Using two criteria (eigenvalues > 1 and scree plots), a 3 factor model was chosen as the most appropriate for the data. The 3 factor model was analyzed to determine which symptoms loaded with which factor. Seven of the 44 items had loadings that were similar across more than 1 factor (3 had similar loadings for all 3 factors). For these 7 items, the highest factor loading was used to determine which factor to assign the item. The three psychosocial factors are described in Table 2, where they are assigned the descriptive names of 1) depression symptoms; 2) anxiety symptoms; and 3) somatic symptoms. These names are descriptive only (as opposed to DSM diagnostic categories) and represent a general description of the symptoms for each factor.

Scales to assess functional impairment were developed based on function questions in two sections. The first section includes gender-specific sets of 13 questions on the degree of difficulty doing a specific activity or task compared to others the same sex and age. These activities/tasks were taken directly from the qualitative study, in which men and women were asked to indicate what are the important activities and tasks that their own sex regularly does to care for themselves, their families, and their communities. Items that were frequently mentioned by the qualitative study respondents were included in the scales. The differences in the content of the sex-specific questions reflects the different responsibilities of each sex, as represented in the qualitative data.

The second component of the function assessment section is an adaptation of the 12-item version of the WHO Disability Assessment Scale (WHO-DAS II). This is a widely used brief assessment of overall functioning that includes ability to do individual activities and tasks, as well as general questions about overall function and ability to work/do normal activities. Unlike the first section, it is not gender specific but is used for all interviewees. As developed, the WHO-DAS asks the respondent to indicate their difficulties *because of health conditions*. For this study, we removed the reference to *health conditions* and asked the respondents to indicate their difficulties due to any cause, because our interest is in improving functionality among program participants regardless of the cause of the dysfunction. The inclusion of the WHO-DAS was done to expand the measurement of functioning beyond only those tasks identified by community members to include items that are considered more universally associated with function and disability.

Descriptions of the scales developed from the instrument (both psychosocial symptoms and function activities) are described in Table 2.

<u>Reliability and Validity Study</u>

The purpose of the reliability and validity study was to determine if the screening instrument could accurately assess the presence and severity of significant psychosocial problems among the target population. Reliability and validity testing included assessment of the following instrument characteristics:
1. Test retest reliability
2. Internal consistency reliability
3. Criterion validity

To ensure the accuracy of data recording, all data from this study was double entered, with all subsequent data analysis conducted using Stata statistical software.

Reliability Testing Methods

Evaluation of test-retest reliability was done by having each interviewer re-interview at least one of their respondents 1-3 days after their first interview. Finding willing respondents for the re-interview proved more difficult than for the initial interview as many were away from their homes for work purposes or did not want to sit through the interview a second time. Scores for each of the subscales were compared at the two time points using pearson correlation coefficients. Analysis of internal consistency was done using Cronbach's alpha calculations.

Validity Testing Methods

Criterion validity was assessed by comparing the scale scores of adults identified as having at least one of the locally-identified psychosocial problems of *fear* and *thinking too much* by themselves and local community leaders ('cases') with the scores of those adults said to have neither of these problems according to themselves and local community leaders ('non-cases').

A team of 14 trained local interviewers and 3 ICMC supervisors carried out data collection with field supervision from JHU faculty. RATA counselors first visited with local knowledgeable people (often community leaders and religious leaders) and asked for a list of adults who had at least one of the two locally-defined psychosocial problems of *fear* and *thinking too much*. They were not asked to distinguish between the different problems, or to name which one they had, but simply whether someone had either one. They were also asked for a list of adults who clearly did not have either of these problems. In this way, we developed 2 lists of local adults: one list for those said by a knowledgeable person to have at least one of the local problems and a list of adults said to have neither of these problems. We did not require the key informants to differentiate between the two types of problems because in the initial qualitative assessment there was not a good differentiation between them. Based on the results of the qualitative study, there did not appear to be a specific local term or syndrome that we could use to define people with psychosocial problems. However, both of these problems, *fear* and *thinking too much*, were often mentioned as important and prevalent problems. It was unclear, however, if people would be able to easily differentiate between people having these problems as the symptoms associated with them were overlapping. That is, someone who has fear also had the problem of thinking too much and vice-versa. Given the lack of differentiation, it was decided to use these two primary

problems to introduce the ideas to the key informants to help explain that we were looking for people who had these types of problems.

The RATA counselors then provided each interviewer with the names of adults to be interviewed in each village, without the designation of whether or not the person was identified as having or not having fear or thinking too much. The interviewers visited the adults they were assigned to and administered the full instrument, at the end asking whether they believed they had each of the locally-described problems in question. When the interview forms were returned the study coordinators and RATA counselors reviewed each form. Where the community referral and the respondent themselves agreed on the presence of a problem, the adult was then regarded as having a psychosocial problem, defined as being a 'case'. Where community referral and respondent agreed that the respondent did not have either of the locally-described problems, that person was considered to have neither of these problems and was defined as a 'non-case'.

Baseline Screening

After completion of the validity study the resulting data suggested that the instrument was appropriate for local use (See Results). Therefore, interviewers began to use the instrument to screen people in the 6 villages for participation in the evaluation study. Those with high enough scores were considered eligible for the study and invited to participate (see Results for cut-off scores). The goal was to enrol a total sample of 400 adults into the intervention study (200 in each study arm: control vs. intervention).

The validity study data suggested that those identified as cases by local people had higher problem scores than non-cases. Therefore, to increase the likelihood of finding eligible persons, interviewers identified persons to interview by knowledgeable local people to name people they thought had the problems of *fear* or *too many thoughts*. Similarly, interviewers asked the respondents themselves if they knew of others who also had these problems.

Table 2: Description of scales from the psychosocial and functional impairment assessment instrument

Depression Symptoms	Anxiety Symptoms	Somatic Symptoms	Male Local Functions	Female Local Functions	WHO DAS Functions
- blaming yourself* - Crying* - don't care about family* - feeling hopeless about the future* - feeling sad* - feeling lonely* - thoughts of ending your life* - feeling of being trapped* - feelings no interest* - difficult to do anything* - feeling of worthlessness* - spacing out - easily angered - don't have direction - everything done goes wrong - stress - can't let the voice out when speak - pale - not wanting to talk	- suddenly scared for no reason[†] - fearful[†] - nervousness or shakiness[†] - heart pounding[†] - trembling[†] - busy by own self (panic) [†] - can't sit, can't stand (feeling restless)[†] - closed breath/difficulty breathing - feeling difficult when having many thoughts* - difficult heart - chaotic thoughts/confusion - trauma - many thoughts	- headache[†‡] - dizziness[†‡] - pain in chest[‡] - pain in lower back[‡] - soreness of muscles[‡] - numbness in parts of your body[‡] - weakness in your body[‡] - feeling tense[†] - no energy* - loss of appetite* - hot body - when sleep, can't sleep well*	- shaving - brushing teeth - brushing hair - go to field/plantation - care of animals - care of children - shopping (buy some rice, fish) - community self-help groups - attend community meeting - attend parties - praying - reciting Koran - earn money - Go to work	- brushing teeth - brushing hair - putting on make-up - cooking - cleaning house - getting water - go to field/rice paddy - gather fire wood - washing clothes - community work - participating in family welfare program - attend parties - praying - reciting Koran - earn money - Go to work	- standing for long periods of time - taking care of your household responsibilities - learning a new task, for example, how to get to a new place - how much of a problem did you have in joining in community activities (for example: festivities/religious activity) in the same way as anyone else can - are you able to do an activity for 10 minutes - go for a long distance by foot - washing your whole body - when put clothes/dress on - when dealing/meeting with people you do not know - maintaining brotherhood with other people - your daily work

* Items from the HSCL depression scale

[†] Items from the HSCL anxiety scale

[‡] Items from the WHO somatic scale

RESULTS

Description of Study Sample

For the validity study, 179 individuals were interviewed. Of these, 106 (59%) were identified by these leaders as having at least one of the two specific problems (*fear* and *thinking too much*) and the remaining 73 (41%) were identified as having neither of these problems.

After completion of the validity study an additional 413 interviews were conducted to screen persons for the intervention study. Therefore, in total, 592 interviews were conducted over the course of 10 days in the 6 study villages. As described in Methods, recruitment for the screening interviews was done using a modified snowballing technique, where the initial interviews were conducted with people identified by local community and religious figures (see description of validity methods above) and each of these initial interviewees was asked if they could refer the interviewers to other people in the community who they thought had these problems. In addition, other community members (exact numbers were not recorded) who were in the area during the interviewing period and requested the opportunity to be interviewed after the purpose of the study was described to them were included. Therefore, while suitable for instrument testing purposes, this sample does not represent a randomly selected, representative sample of the community and the results are not generalizeable to the total adult population in the area. Table 3 presents a basic description of the study sample.

Table 3: Study Sample Description

	Total Sample (N=592)	Validity Sample (N=179)
Sex		
Male, N (%)	276 (47)	113 (63)
Female, N (%)	316 (53)	66 (37)
Age		
Less than 30 years, N (%)	64 (11)	13 (7)
30-49 years, N (%)	266 (45)	91 (51)
50-69 years, N (%)	203 (34)	60 (34)
70 or more years, N (%)	59 (10)	15 (8)
Marital Status		
Single, N (%)	34 (6)	12 (7)
Married, N (%)	470 (79)	144 (80)
Widow/Widower, N (%)	88 (15)	23 (13)
Psychosocial Scales[1]		
Depression symptoms (19 items), Mean (SD)	14.2 (10.6)	14.1 (11.3)
General anxiety symptoms (13 items), Mean (SD)	18.6 (10.0)	18.3 (10.0)
Somatic anxiety symptoms (12 items), Mean (SD)	21.0 (8.3)	19.9 (8.8)
Total symptoms[2] (44 items), Mean (SD)	53.8 (25.4)	52.3 (26.7)
Functional Impairment Scales[3]		
Local functions, male (14 items), Mean (SD)	9.9 (9.2)	9.3 (9.2)
Local functions, female (16 items), Mean (SD)	11.5 (9.8)	12.4 (10.2)
WHO DAS items (11 items), Mean (SD)	9.3 (6.8)	8.6 (7.1)

1. Psychosocial scales are made up of items scored on a 4-point scale: 0=not experiencing the symptom at all, 1=experiencing it rarely, 2=sometimes, 3=often.
2. Total symptoms includes all 44 symptoms.
3. Functional impairment scales are made up of items scored on a 5-point scale: 0=no problem at all doing the activity/task, 1=a little difficulty, 2=a moderate amount, 3=a lot, 4=cannot do.

While the validity study was conducted separately from the screening interviews, some of the validity study analyses can be done with all 592 interviews, not just the 179 from the original study. Since using all interview data provides more accurate results we have reported the results of analysis for all interviews where appropriate. For other analyses (the case vs noncase data) only the 179 validity study interviews are appropriate and therefore only those were used. Whether the full sample or only the validity sample were used in an analysis is stated for each table of results.

Instrument Reliability

Table 4 shows test-retest reliability analysis results, based on the 17 respondents who were re-interviewed by the same interviewers 1-3 days after the first interview. Normally, at least 30 interviews would be repeated, however very few interviewees were willing to be re-interviewed. To gather additional information about the test-retest reliability, additional interviews will be conducted as part of the post-intervention follow-up survey. Test-retest reliability is assessed using the Pearson correlation coefficient, which provides a measure of how similar each subscale score is on the first and second interviews. This in turn provides an indicator of the extent to which the respondents tend to give the same answer to the questions constituting the scale when asked on different occasions and by the same interviewer.

When assessing test-retest reliability, Pearson correlation coefficient scores of .7 are considered to be acceptable. The results in Tables 4 suggest that the scales for the depression and total symptom scores can be reliably be assessed over time. The other symptom scales show borderline reliability. The male function scales and the items derived from the WHO DAS scale performed poorly.

Table 4: Test-retest comparisons

Total Sample	First Interview		Repeat Interview		Correlation[1]
	N	Mean (sd)	N	Mean (sd)	
Psychosocial Symptom Scores					
Depression symptoms score	17	19.4 (14.3)	17	18.6 (15.3)	.96
General anxiety score	17	24.9 (7.9)	17	21.2 (9.8)	.65
Somatic anxiety score	17	24.5 (6.1)	17	24.2 (7.1)	.60
Total symptom scale score[2]	17	68.8 (24.1)	17	64.0 (28.7)	0.90
Functional Impairment Scores					
Male local functions	8	21.4 (16.8)	8	17.3 (13.6)	0.32
Female local functions	9	12.6 (7.6)	9	10.6 (8.0)	0.64
WHO DAS scale	17	13.6 (9.3)	17	11.6 (6.6)	0.35

1. Pearson's correlation coefficient
2. Total symptoms scale includes the three syndrome scales plus the additional local qualitative symptoms

Internal consistency reliability measures the extent to which questions that assess the same underlying concept agree or disagree. If these questions disagree this suggests that either the questions themselves are unreliable, or they are not really measuring the same concept. As described in Appendix C, internal consistency reliability is measured using Cronbach's alpha. Scores should be at least .7 and ideally >.8.

Table 5 shows the Cronbach's alpha scores on each of the scales for the total sample and separately by sex. Alpha scores for the somatic scale is acceptable and for all other scales are very good.

Table 5: Cronbach's alpha reliability scores

	Total Sample (N=592)	Men (N=276)	Women (N=316)
Psychosocial Symptom Scores			
Depression symptoms score	0.88	0.88	0.89
General anxiety score	0.89	0.89	0.90
Somatic anxiety score	0.84	0.86	0.83
Total symptom scale score[1]	0.94	0.94	0.94
Functional Impairment Scores			
Male local functions		0.90	
Female local functions			0.90
WHO DAS scale	0.84	0.86	0.82

1. Total symptoms scale includes all 44 symptoms

Instrument Validity

Community and religious leaders provided 103 names of people they identified as having the problems of *fear* and/or *thinking too much* and 86 (83%) of the respondents agreed (defined as 'cases'). The community and religious leaders also provided the names of 79 people they identified as having neither of these types of problems, of whom 23 (30%) agreed with this designation (defined as 'non-cases').

Table 6 examines the criterion validity of the scales by comparing scale scores of the subsample of 109 adults defined as likely 'cases' and 'non-cases' based on the assessments by the local community and religious leaders and self-assessment by the respondents themselves (See Methods).

For all 109 adults, differences in scores between 'cases' and 'non-cases' were statistically significant on all scales, except for the female local functional impairment scale. Differences tended to be in the direction expected: higher scores among 'cases' and the absolute difference in most scale scores between 'cases' and 'non-cases' were large, with the 'cases' having scores

more than twice as high as the 'non-cases', except for the female functional impairment score.

Table 6: Comparison of scale scores for cases versus non-cases

	Score Range (Min, Max)	Cases[1] (N=86)	Non-Cases (N=23)	Difference (p-value)[2]
Psychosocial Symptom Scores				
Depression symptom score, mean (se)	0, 49	16.8 (1.2)	5.6 (1.0)	< 0.0001
General anxiety score, mean (se)	0, 39	22.4 (1.0)	6.7 (1.5)	< 0.0001
Somatic anxiety score, mean (se)	0, 36	22.9 (0.8)	11.0 (1.8)	< 0.0001
Total symptom scale score[3], mean (se)	0, 132	62.1 (2.5)	23.3 (3.8)	< 0.0001
Functional Impairment Scores				
Male local functions[4], mean (se)	0, 56	10.9 (1.5)	4.6 (1.3)	0.02
Female local functions[5], mean (se)	0, 64	11.3 (1.5)	X	X
WHO DAS scale, mean (se)	0, 44	9.4 (0.7)	4.5 (1.0)	0.002

1. Cases include those who were identified by community referral as having either of the two local problems and those who self-identified as having either, or both, of the two local problems.
2. P-value for the statistical significance of the difference in scale scores by caseness.
3. Total symptoms scale includes all 44 symptoms
4. Sample size for the male specific scale are as follows: male cases (n=47), male non-cases (n=17).
5. Comparisons for statistical significance were not made for the female function scale as the number of non-cases was too small (n=6) to make an appropriate comparison.

Screening into intervention evaluation study

Eligibility for inclusion in the counseling program is based on severity of psychosocial symptoms and associated functional impairment. The cutoff score which was used in the screening interviews to determine eligibility for inclusion in the intervention study was determined based on data collected during the validity study described above. During the validity study, we found that the 'cases' were highly symptomatic on the total psychosocial problems scale score, with a mean score on the total symptoms scale of 62 points (sd 24). Since choosing the mean score of the cases as the cut-off would limit inclusion to only half of all cases (the more severe half) we chose a cut-off score of one standard deviation below this mean (cutoff = 38 points) in order to include the bulk of persons who would be considered as cases. This score of 38 points gave good discrimination from non-cases, who mean was > 10 points less (23.3 points).

When reviewing interviews for those considered eligible for the intervention study we also

reviewed the validity study interviews. Therefore, the interviews of a total of 592 adults were screened to evaluate their eligibility for inclusion in the study. The 592 included all of the participants from validity study (n=179) along with the additional 413 who were assessed subsequent to the validity study. Using the inclusion criteria described above, 421 adults met the cut-off criteria of a total psychosocial problems score greater than 38 points and had some degree of functional impairment and were therefore included in the evaluation study. This number included 129 (72%) of the validity sample. Five additional people were also included because their scale scores were between 37-37.7 and they had high functional impairment scores. This gave the total study sample eligible for inclusion as N=426, or 72% of the total screened sample.

Table 7 presents the distribution of these 426 respondents who met study inclusion criteria described above, across the 6 villages and by gender. Although there is some variation in the proportion of men and women in the different villages meeting eligibility criteria, across intervention and control status, the numbers were not statistically significantly different (p=.07) though the proportion of females in the control villages was higher than in the intervention villages (Table 8). Table 8 presents general information about the sample that met eligibility criteria for the evaluation study.

Table 7: Study Sample That Met Eligibility Criteria (N=426)

Village	Blang Gandai	Blang Ketumba		Batee Raya	Abeuk Usong		Abeuk Budi	Peruaden
Intervention Status[*]	I	C		I	C		I	C
Study Size	72	62		72	76		74	70
Males (%)	32 (44)	30 (48)		37 (51)	31 (41)		40 (54)	25 (36)
Females (%)	40 (56)	32 (52)		25 (49)	45 (59)		34 (46)	45 (64)

Table 8: Baseline scale scores for evaluation sample who met study eligibility criteria

	Total Study Sample (N=426)	Intervention Sample (N=218)	Control Sample (N=208)	p-value[1]
Sex				
Male, N (%)	195 (46)	109 (50)	86 (41)	
Female, N (%)	231 (54)	109 (50)	122 (59)	.07
Age				
Less than 30 years, N (%)	32 (8)	18 (8)	14 (7)	
30-49 years, N (%)	191 (45)	93 (43)	98 (47)	
50-69 years, N (%)	158 (37)	80 (37)	78 (37)	
70 or more years, N (%)	45 (11)	27 (12)	18 (9)	.53
Marital Status				
Single, N (%)	18 (4)	12 (5)	6 (3)	
Married, N (%)	335 (79)	172 (79)	163 (78)	
Widow/Widower, N (%)	73 (17)	34 (16)	39 (19)	.20
Psychosocial Symptoms Scales[2]				
Depression symptoms, Mean (SD)	17.9 (10.1)	17.9 (10.9)	17.9 (9.2)	.99
General anxiety symptoms, Mean (SD)	22.9 (8.1)	23.4 (8.3)	22.4 (7.8)	.21
Somatic anxiety symptoms, Mean (SD)	24.7 (5.8)	24.5 (6.2)	25.0 (5.4)	.43
Total symptoms[3], Mean (SD)	65.5 (19.2)	65.8 (20.5)	65.2 (17.8)	.77
Functional Impairment Scales[4]				
Local functions, male, Mean (SD)	12.0 (9.7)	10.8 (9.0)	13.5 (10.5)	.05
Local functions, female, Mean (SD)	13.3 (10.2)	11.5 (10.5)	14.9 (9.7)	.01
WHO DAS items, Mean (SD)	11.0 (6.9)	10.2 (7.1)	12.0 (6.7)	.01

1. p-value refers to the likelihood of significant differences between the intervention and control samples. A p-value > .05 indicates the two samples are not statistically significantly different on that characteristic.
2. Psychosocial scales are made up of items scored on a 4-point scale: 0=not experiencing the symptom at all, 1=experiencing it rarely, 2=sometimes, 3=often.
3. Total symptoms includes all 44 symptoms
4. Functional impairment scales are made up of items scored on a 5-point scale: 0=no problem at all doing the activity/task, 1=a little difficulty, 2=a moderate amount, 3=a lot, 4=cannot do.

After all interviews were completed, the interviewers returned to the study villages. In the three control villages they personally told each respondent whether or not they met criteria for the program and if they did, that they would receive the intervention in the second phase, beginning in January. In the intervention villages the interviewers informed all of the respondents who did not meet eligibility criteria of their non-participation status. The RATA counselors were provided with the names and locations of the participants in the intervention villages who met inclusion criteria and visited each one to invite them to participate in the group counseling program.

DISCUSSION AND CONCLUSIONS

Reliability and validity testing showed a distinct pattern in the performance of the various scales. On both measures of reliability (test-retest and internal consistency) the symptom-based scales show good psychometric properties in this population. However, the function scales performed poorly in the test-retest analysis, particularly for the WHO DAS scale and the male-specific

functioning scale. The low correlation for the WHO DAS scale was primarily due to the male respondents as when the analysis was done separately by sex, the correlation among females was 0.63 and among males was 0.20. One possibility for the lack of consistency is that some of the items may not be regularly done (i.e. community work, attending parties) so that when asked to describe the amount of difficulty doing these more 'rare' activities, recall could have differed with time. Another possibility is that some factor(s) that we did not measure were associated with a true shift in the perceptions of difficulty in engaging the different tasks. We had several refusals to be interviewed for the second interview, which could be associated with a self-selection bias which may then be associated with changes in response patterns. Of all types of reliability, in general test-retest is often the weakest as it relies on non-change in perceptions over a short period of time. The small sample size for the current test-retest analysis makes it difficult to make comprehensive conclusions and thus, rather than remove these questions, we will continue to use them and conduct additional repeat interviews during the post-intervention assessment for further evaluation.

Tests of criterion validity found that the symptom scale scores consistently matched the criterion chosen for this study: a local community or religious leader's identification along with self-identification of having either the problem of *fear* or *thinking too much*: The respondents identified by both as having either of the problems ("cases") showed significantly higher levels of symptoms than those identified as neither of these problems ("non-cases"). The exception to this was the female function score. The analysis of this scale resulted in counter-intuitive results: those identified as "non-cases" had higher dysfunction scores than those identified as "cases". This is primarily due to a single outlier among the female "non-cases" who had a functional impairment score of 43. If that respondent was removed from the comparison, the mean score of the "non-cases" would be 9.4 (se: 1.8). This average score is lower, but given the small sample size it is difficult to say anything meaningful about statistical significance.

In addition to the main purpose of the study - developing an acceptable, reliable, and valid instrument – we continued with the interviewing process and screened an additional 592 adults for inclusion in the intervention evaluation study. A total of 426 met the study inclusion criteria and were invited to either participate in the group counseling program (intervention villages) or were told that their villages would be receiving the program in the second stage and were invited to be re-screened prior to that (in January).

The interviewers returned to the control villages and informed them that their villages would receive the program in the second stage. For those who met the program criteria, they were asked if they agreed to be revisited in January and to participate in the program then. Thirteen of the 208 respondents indicated they did not want to be revisited or participate. In addition, several of the villagers who were told that they did not meet inclusion criteria (it was explained that they did not have the types of problems the program is good for) still wanted the opportunity to participate. It was decided that anyone who wanted to would be revisited in January and re-assessed for inclusion in the second stage of the program.

Visiting the intervention villages and informing the screening respondents whether they met the study inclusion criteria or not was completed after the JHU team left the area. The interviewers were going to meet with the villagers who did not meet criteria and explain their non-

participation to them and the counselors were going to meet with those who met criteria and invite them to participate. During this initial meeting, the counselors would also begin to form the counseling groups.

Since the completion of this field study, the intervention has begun, with field supervision being conducted by ICMC staff. The 8-weeks of counseling is expected to be completed by the end of November and will be followed by a brief qualitative study to identify unexpected (positive and negative) outcomes of the intervention. This will be followed in January by a formal quantitative follow-up assessment in January of all study eligible adults.

RECOMMENDATIONS

The primary recommendation would be to continue to use this instrument development and validation process (both the qualitative and quantitative components) to improve program development, monitoring and evaluation. This evaluation study represents the full implementation of the process, which can be replicated with other ICMC programs. The methods are particularly useful in situations where need has not been well characterized and where the impact of interventions has not been demonstrated.

In addition, further investigation needs to be done to understand the low reliability of the male function items. This will be done first by conducting additional repeat interviews to bolster the sample size for the test-retest analysis. In addition, when responses for the functions differ between the two time points, this will be explored further with the respondents.

Finally, to finalize the validity analysis, we will identify additional women without either of the two local problems of *fear* and *thinking too much* to assess. These data will supplement the small number (n=6) of women thus identified in this initial study.

APPENDICES

Appendix A: Example of Village Socialization Process

<div align="center">

Visiting Pak Keucik (head of the village)

</div>

Two counselors visited Pak Keucik (head of the village) and talked about Rata and the program. "Rata is the non-governmental organization that helps the victims of violence and his/her family. We had been worked in Banda Aceh, Meulaboh, Langsa, Bireuen and Pidie. Our office is in Geulanggang Tengeh village. Now, Rata want to help people in this village who have symptoms like saket ulee, mumang, teutahe tahe, trauma and other similar symptoms. For them, we want to create kelompok peuga peuga haba to reduce their beban hate. Since we will work in this village, could us have a community meeting to inform the villagers about this?" Pak keucik needed to discuss about this in the village meeting and He asked the counselors to come back in few days. The counselors went back to pak keucik to ask about the schedule of the community meeting.

<div align="center">

The first community meeting

</div>

All counselors (8 counselors) came

Greeting and Introduction

Thank you for your attendance to be in this meeting today. First, we will introduce our selves. My name is ____, ____, ____, ____."

Rata

"We are from a non governmental organization. Its name is RATA. Rata stands for Rehabilitation Action for Torture victims in Aceh. RATA has assisted the community since 1999 in some areas, such as Banda Aceh, Meulaboh, Langsa, Pidie and Bireuen."

Program

"Rata helps conflict, tsunami, torture and disaster victims. Our assistance is for handle problem of beban hate/heart burden by being a friend to talk to and to discuss with."

Musibah/disaster (with showing pictures of some bad situations in man life – disaster, conflict, a wife with a sick husband)

" In our life, we often experience some difficult situation or disaster, such as tsunami, earthquake, conflict or our husband or other member of our family get sick. What is the difficult situation or disaster that people in this village ever experienced ?"

Symptoms (with showing pictures of some symptoms – crying, spacing out, having many thoughts, fear, headache)

"Then, after getting through this difficult situations and disaster, it's a natural reaction that people having heart burden and becoming often crying, having many thoughts, sad, fear, headache and trauma. Is there anything else alike that caused by the heart burden after the conflict that people in this village still feels?"

Coping (with showing pictures of some coping mechanisms – praying, going to the field, chatting)

"Usually to handle burden of heart, people praying, go to work in the field, talk to others. What are the activities that people in this village usually do to reduce burden of heart ?"

Kelompok Peuga-peuga haba (with showing pictures of kelompok peuga peuga haba)
"Usually, if we had heart burden, if we keep it by ourselves, then it would be felt so heavy. It also happends if we are having problems, if we are thinking it by ourselves, it feels so difficult to find solution. But if we share our heart burden, then it will be lighter and if we share our problem and discuss our problems together, we will find the solution."

"Before coming to this village, we had also assisted other villages in Bireuen. In those villages, we were seeing people who have many thoughts, fear, sadness, headache, heart pounding and spacing out."

"To help them reducing their heart burden, we gave assistance by making kelompok peuga peuga haba."

"In kelompok peuga peuga haba, people can share their heart burden and many thoughts and together find the solution how to deal with daily problems. This is the program that we also will do in this village."

"We need to explain from this beginning that Rata's assistance is not giving any money, materials or medicines, but to help reducing the heart's burden by becoming a friend to talk to and discuss."

Process toward intervention in this village
"On the August 14th until 10 days after that, they will be people from Rata who will interview all the villagers, especially people who has symptoms like mumang, teutahe tahe, trauma, lee pikiran. The result of interviews will show who are people in this village that can be helped with our program. We will do all this interview in 6 villages, which are Abuk Usong, Abuk Budi, Batee Raya, Blang Gandai, Blang Ketumba and Peuraden. Since we are not having human resource to serve all these villages at once, we will serve 3 villages first and after that we will continue to serve other 3 villages. We will choose the 3 villages randomly." After all those explanation, we asked whether any questions from the villagers. These are examples of some of the questions from the villagers :
1. When Rata will give service in this village?
2. Will Rata give medicines?
3. So, kelompok peuga peuga haba just will talk?
4. If someone is having stress, can you give a free service (including transportation and meals)?

The closing

"Ladies and gentleman, we hope that our meeting is a pleasure for all of us. Before closing this meeting today, we would like to know whether it's possible to have this kind of meeting again in 1 or 2 weeks ahead?"

The Second Community Meeting: All counselors (8 counselors) came
Greeting and Introduction

Thank you for your attendance to be in this meeting today. First, we will introduce our selves. My name is ____, ____, ____, ____."

RATA

"We are from a non governmental organization. Its name is RATA. Rata stands for Rehabilitation Action for Torture victims in Aceh. RATA has assisted the community since 1999 in some areas, such as Banda Aceh, Meulaboh, Langsa, Pidie and Bireuen."

Review about First Community Meeting

"As we had explained on the 1st meeting before, the assistance that RATA will give in this village is to help people having heart burden (such as : sadness, difficult heart, fear, anger, headache, heart pounding, spacing out, etc) by making kelompok peuga peuga haba."

"In kelompok peuga peuga haba, we will share heart burden and many thoughts, we can share heart burden and many thoughts and together looking for solutions for daily problems."
"We need to explain from this beginning that Rata's assistance is not giving any money, materials or medicines, but to help reducing the heart's burden by becoming a friend to talk to and discuss."

Process toward intervention in this village
"On the August 14th until 10 days after that, they will be people from Rata who will interview all the villagers, especially people who has symptoms like mumang, teutahe tahe, trauma, lee pikiran. The result of interviews will show who are people in this village that can be helped with our program. We will do all this interview in 6 villages, which are Abuk Usong, Abuk Budi, Batee Raya, Blang Gandai, Blang Ketumba and Peuraden. Since we are not having human resource to serve all these villages at once, we will serve 3 villages first and after that we will continue to serve other 3 villages. We will choose the 3 villages randomly."

Film

"We will play short video clip about kelompok peuga peuga haba that had been done by Rata in some villages in Bireuen."

Discussion
"Are there any question about Rata and its program ?"
Examples of questions :
1. When Rata will give service in this village ?
2. Can kelompok peuga peuga haba cure the physical sickness ?
3. In kelompok peuga peuga haba, how many meeting in a week and when will be the time ?

Small Competition
"After hearing all the explanation from Rata, we will do a small fun activity. So, let's spilt up into two big groups. We will give a question to each group then to other group repeatedly."

Questions :
1. What is the advantage of sharing to others ?
2. What kind of people can be helped by Rata ?
3. What is the purpose of Rata's program ?
4. What is heart burden ?
5. What does Rata stands for ?
6. What people usually do to reduce heart burden ?
7. What is the purpose of kelompok peuga peuga haba ?
8. What do we feel if sharing our heart burden ?
9. What is the example of heart burden for people in this village ?
10.What will be happen if we keep the problem by own self ?

The Closing

"Ladies and gentleman, we hope that our meeting is a pleasure for all of us. We will end this meeting

with praying."

Informal Socializations

This activity had been done by 2-4 counselors at a time. They come to the village together and after they have arrived in the village, they will split up and go to place where they see people is gathering. They are sitting around at about 30 minutes in each place.

Activities :
1. Sitting in the coffee shop
2. Sitting in the shops
3. Sitting in villager's house

Pointers of discussion :
Rata and its program

"We are from Rata, a non governmental organization. Our office is in Geulanggang Tengeh. RATA will give assistance in this village to help people having heart burden (such as : sadness, difficult heart, fear, anger, headache, heart pounding, spacing out, etc) by making kelompok peuga peuga haba."

"In kelompok peuga peuga haba, we will share heart burden and many thoughts, we can share heart burden and many thoughts and together looking for solutions for daily problems."
"We need to explain from this beginning that Rata's assistance is not giving any money, materials or medicines, but to help reducing the heart's burden by becoming a friend to talk to and discuss."

Daily lives in the village
"Mam/Sir, what people in this village do daily?"

Male and female (villagers) activities
"Mam/Sir, what are the activities that women and man usually do together in this village?"

Looking for names of people having symptoms and not having symptoms in the village
"Mam/Sir, could you please tell me the names of villagers here that you know seems to have many thoughts, sadness, headache, spacing out or trauma?"
"Mam/Sir, could you please tell me the names of villagers here that you know seems not having many thoughts, sadness, headache, spacing out or trauma?"

Process toward intervention in this village
"On the August 14th until 10 days after that, they will be people from Rata who will interview all the villagers, especially people who has symptoms like mumang, teutahe tahe, trauma, lee pikiran. The result of interviews will show who are people in this village that can be helped with our program. We will do all this interview in 6 villages, which are Abuk Usong, Abuk Budi, Batee Raya, Blang Gandai, Blang Ketumba and Peuraden. Since we are not having human resource to serve all these villages at once, we will serve 3 villages first and after that we will continue to serve other 3 villages. We will choose the 3 villages randomly."

This informal socializations was being done 3 times per village.

Appendix B: Version of the instrument used in Validity Study and Baseline Screening

Part A: Assessment of Function

I am going to read a list of tasks and activities. These are tasks and activities that other people around here told us were important for men/women (refer to sex of the respondent) to be able to do. For each task I am going to ask you how much more difficulty you are having doing it compared with WHAT YOU THINK OTHERS WHO ARE ABOUT YOUR SAME AGE AND SEX NORMALLY DO. You should tell me whether you are having no more difficulty, a little more, a moderate amount more, or a lot more, or you often cannot do that task.

To make it easier to remember I have a card here with pictures. Each picture represents a different amount of difficulty. Show the respondent the card illustrating levels of difficulty. Point to each picture as you describe it.

The first picture shows someone who has no more difficulty than most other men/women of your age. The second picture shows someone who has a little more difficulty. The third picture shows someone who is having a moderate amount more difficulty. The fourth picture shows someone who is having a lot more difficulty and the last shows someone who is having so much difficulty they often cannot do the task. For each task or duty, I will ask you to point to the picture which shows how much difficulty you are having in doing that task, compared with what you think others who are about your same age and sex normally do.

Now say each task, and after each one say: In the past two weeks are you having no more difficulty, a little more, a moderate amount more, a lot more, or are having so much difficulty that you often cannot do the task? pointing to each picture as you say it. Record the response by marking the appropriate box next to the activity or task in the table below.

Before each item, say to the person 'In the last two weeks, how much difficulty have you had with…"

Male	In the last two weeks, amount of difficulty doing each activity					
	No difficulty	a little difficulty	A moderate amount	A lot of difficulty	So much, cannot do it	not relevant for me
A01 shaving (Q)	0	1	2	3	4	9
A02 brushing teeth (Q)	0	1	2	3	4	9
A03 brushing hair (Q)	0	1	2	3	4	9
A04 go to field/plantation (Q)	0	1	2	3	4	9
A05 care of animals (Q)	0	1	2	3	4	9
A06 care of children (Q)	0	1	2	3	4	9
A07 shopping (buy some rice, fish) (Q)	0	1	2	3	4	9
A08 community self-help groups (Q)	0	1	2	3	4	9
A09 attend community meeting (Q)	0	1	2	3	4	9
A10 attend parties (Q)	0	1	2	3	4	9
A11 praying (Q)	0	1	2	3	4	9
A12 reciting Koran (Q)	0	1	2	3	4	9
A13 *earn money*	0	1	2	3	4	9
A14 Go to work	0	1	2	3	4	9

Before each item, say to the person 'In the last two weeks, how much difficulty have you had with…"

Female	In the last two weeks, amount of difficulty doing each activity					
	No difficulty	a little difficulty	A moderate amount	A lot of difficulty	So much, cannot do it	not relevant for me
A15 brushing teeth (Q)	0	1	2	3	4	9
A16 brushing hair (Q)	0	1	2	3	4	9
A17 putting on make-up (Q)	0	1	2	3	4	9
A18 cooking (Q)	0	1	2	3	4	9
A19 cleaning house (Q)	0	1	2	3	4	9

Respondent ID# _____

A20 getting water (Q)	0	1	2	3	4	9
A21 go to field/rice paddy (Q)	0	1	2	3	4	9
A22 gather fire wood (Q)	0	1	2	3	4	9
A23 washing clothes (Q)	0	1	2	3	4	9
A24 community work (Q)	0	1	2	3	4	9
A25 participating in family welfare program (Q)	0	1	2	3	4	9
A26 attend parties (Q)	0	1	2	3	4	9
A27 praying (Q)	0	1	2	3	4	9
A28 reciting Koran (Q)	0	1	2	3	4	9
A29 earn money	0	1	2	3	4	9
A30 Go to work	0	1	2	3	4	9

Respondent ID# _____

Before each item, say to the person 'In the last two weeks, how much difficulty have you had with…"

male and female	In the last two weeks, amount of difficulty doing each activity					
	No difficulty	a little difficulty	A moderate amount	A lot of difficulty	So much, cannot do it	not relevant for me
A31. standing for long periods of time	0	1	2	3	4	9
A32. taking care of your household responsibilities	0	1	2	3	4	9
A33. learning a new task, for example, how to get to a new place	0	1	2	3	4	9
A34. how much of a problem did you have in joining in community activities (for example: festivities/religious activity) in the same way as anyone else can	0	1	2	3	4	9
A35. how much does your health problems influence your feeling?	0	1	2	3	4	9
A36. (are you able to do an activity for 10 minutes)	0	1	2	3	4	9
A37. go for a long distance by foot	0	1	2	3	4	9
A38. washing your whole body (Q)	0	1	2	3	4	9
A39. when put clothes/dress on	0	1	2	3	4	9

Respondent ID# _____

Before each item, say to the person 'In the last two weeks, how much difficulty have you had with…"

male and female	In the last two weeks, amount of difficulty doing each activity					
	No difficulty	a little difficulty	A moderate amount	A lot of difficulty	So much, cannot do it	not relevant for me
A40. when dealing/meeting with people you do not know	0	1	2	3	4	9
A41. maintaining brotherhood with other people	0	1	2	3	4	9
A42. your daily work	0	1	2	3	4	9

	Not at	a little	A moderate	A lot	extremely
A43. From things that we had discussed before, how much have they been disturbing your life?	0	1	2	3	4

A44. In the past 30 days, for how many days were these difficulties present?

A45. In the past 30 days, how many days that you were totally unable to carry out your daily activities or work?

A46. In the past 30 days, how many days you cannot do your daily activities or work fully in a day not including going to prayer on Fridays

<u>Part B</u>- Psychosocial Assessment Instrument

I am going to read you a list of statements. For each one I am going to ask you how much you have felt like that IN THE LAST TWO WEEKS, including today.

Say each statement, and after each one ask how often the respondent has felt like that in the last 2 weeks. Repeat the categories after each statement and let the respondent choose one. Record the response by circling the appropriate box next to the symptom.

Before each item, say to the person "In the last two weeks, how often did you feel …"

Symptoms	Not at all	Rarely	Sometimes	Often
B01. headache (S1/A8/Q1)	0	1	2	3
B02. dizziness (S2/A3/Q2)	0	1	2	3
B03. pain in chest (S3/Q3)	0	1	2	3
B04. pain in lower back (S4)	0	1	2	3
B05. soreness of muscles (S5)	0	1	2	3
B06. numbness in parts of your body (S6)	0	1	2	3
B07. weakness in your body (S7/Q4)	0	1	2	3
B08. suddenly scared for no reason (A1)	0	1	2	3
B09. fearful (A2/Q5)	0	1	2	3
B10. nervousness or shakiness (A4/Q6)	0	1	2	3
B11. heart pounding (A5/Q7)	0	1	2	3
B12. trembling (A6)	0	1	2	3
B13. feeling tense (A7)	0	1	2	3
B14. busy by own self (panic) (A9)	0	1	2	3

Before each item, say to the person "In the last two weeks, how often did you feel …"

Symptoms	Not at all	Rarely	Sometimes	Often
B15. can't sit, can't stand (feeling restless) (A10/Q8)	0	1	2	3
B16. no energy (D1/Q9)	0	1	2	3
B17. blaming yourself (D2)	0	1	2	3
B18. crying (D3/Q10)	0	1	2	3
B19.don't care about family (D4)	0	1	2	3
B20.loss of appetite (D5/Q11)	0	1	2	3
B21.when sleep, can't sleep well (D6/Q12)	0	1	2	3
B22.feeling hopeless about the future (D7)	0	1	2	3
B23.feeling sad (D8/Q13)	0	1	2	3
B24. feeling lonely (D9)	0	1	2	3
B25. thoughts of ending your life (D10)	0	1	2	3
B26. feeling of being trapped (D11)	0	1	2	3
B27. feeling difficult when having many thoughts (D12)	0	1	2	3
B28. feelings no interest (D13)	0	1	2	3
B29. difficult to do anything (D14)	0	1	2	3
B30. feeling of worthlessness (D15)	0	1	2	3
B31. difficult heart (Q14)	0	1	2	3
B32. spacing out (Q15)	0	1	2	3
B33. easily angered (Q16)	0	1	2	3
B34. don't have direction (Q17)	0	1	2	3

Before each item, say to the person "In the last two weeks, how often did you feel …"

Symptoms	Not at all	Rarely	Sometimes	Often
B35. chaotic thoughts/confusion (Q18)	0	1	2	3
B36. everything done goes wrong (Q19)	0	1	2	3
B37. stress (Q20)	0	1	2	3
B38. trauma (Q21)	0	1	2	3
B39. can't let the voice out when speak (Q22)	0	1	2	3
B40. hot body (Q23)	0	1	2	3
B41. pale (Q24)	0	1	2	3
B42. closed breath/difficulty breathing (Q25)	0	1	2	3
B43. not wanting to talk (Q26)	0	1	2	3
B44. many thoughts (Q27)	0	1	2	3

Respondent ID# _____

Part C - Self Assessment

	yes	no
C1. Lately, do you have fear?	1	0
C2. Has anyone said that lately you have fear ?	1	0

If answer to C1 is yes, ask this question:

C3. How long have you had this ? _____days _____ weeks_____months _____year

	Yes	no
C4. Lately, do you have many thoughts?	1	0
C5. Has anyone said that lately you have many thoughts?	1	0

If answer to C4 is yes, ask this question:

C6. How long have you had this ? _____days _____ weeks_____months _____year

	yes	no
C7. Have you ever got beaten up in your head or another head injury?	1	0
C8. If yes, did you faint/loss your consciousness at that time?	1	0

Part D – Coping

For Questions D01a-D09a (first columns): People have told of many different activities people sometimes do to help themselves feel better. I am going to read some of these activities and for each one I am going to ask you how often you do this activity to help yourself when you feel bad.

After all of section A go to Questions D01b-D09b (second columns): I am going to go through the activities that you said you sometimes do when you feel bad. For each one, please tell me how it made you feel: A lot worse, a little worse, no difference, a little better, a lot better.

A.	Not at all	Rarely	Somew hat	Often	**B.**	A lot worse	A little worse	No differ ence	
D01a. Pray	0	1	2	3	D01b. Pray	1	2	3	
D02a. Recite Koran	0	1	2	3	D02b. Recite Koran	1	2	3	
D03a. Earn money	0	1	2	3	D03b. Earn money	1	2	3	
D04a. Sitting together to chat	0	1	2	3	D04b. Sitting together to chat	1	2	3	
D05a. Go (walk) to please own heart	0	1	2	3	D05b. Go (walk) to please own heart	1	2	3	
D06a. Discussion	0	1	2	3	D06b. Discussion	1	2	3	
D07a. Listen to the advice from wise men	0	1	2	3	D07b. Listen to the advice from wise men	1	2	3	
D08a. Go to find recreation for own self	0	1	2	3	D08b. Go to find recreation for own self	1	2	3	
D09a. Play soccer or volley	0	1	2	3	D09b. Play soccer or volley	1	2	3	

Appendix C: Study Explanation Read to Respondents Prior to Interview

Hello, my name is _____. I work for RATA and we are asking people about problems that some people have. We were recently here in your village to introduce our organization and present some of the work that we are going to be doing with people in your village.

Today, and for the next two weeks, we are talking with people in your village about problems that some people might have. To go through all of the questions will take about 40 minutes, do you have the time right now?

If the respondent answers no, thank them. If the respondent answers yes, then explain the following:

During these two weeks we will be asking many of the adults in your village these questions. At the end of the two weeks, we will return to each one of you to talk more about our program and what services we can provide to you. We are not providing financial services and will not be able to provide money or other goods to any of the people in your village.

Ask that the interview be conducted in private. If this is questioned by anyone, explain that this is an important part of our procedure, and that we have found that some people give different answers when there are other people present.

Appendix D: Explanation of Reliability and Validity Concepts

Reliability

Reliability refers to the extent to which different measures of the same concept agree with each other. It can refer to measurements taken at the same time, or different times. To be useful an instrument must have good local reliability, which must therefore be tested whenever a questionnaire is changed (including translation) or used among a new population.

Test-Retest Reliability

Testing reliability over time is an important measure of reliability. This is called test-retest reliability and is implemented by giving the questionnaire to the same subject on two different occasions. It is usually done at least a day later, to reduce the effect of memory on the responses, but not too long because what is being measured may actually change (mood, for example). Therefore, the repeat interview is usually done 1-7 days after the first interview. For this study the second interview was done by the same interviewer, therefore this is not an evaluation of inter-rater reliability (which requires different interviewers administering the questionnaire). Comparison of the results of the first and second interviews is the measure of test-retest reliability. To make this comparison, a summary scale is first created using all the questions on the same topic (in this case each of the psychosocial and functional impairment subscales) and calculated for both the first and second interview. Test-retest reliability is tested by measuring correlations between these scores. Opinions vary as to what is an acceptable score, although correlations above 0.7 are considered desirable for test-retest reliability. A problem arises in interpreting low scores. These may be due to a poor instrument, or to using different interviewers, or because the concept being measured has changed. Partly for these reasons, test-retest, and inter-rater, reliability are not generally considered as important as internal consistency reliability (Streiner et al 1995).

Internal Consistency Reliability

This refers to how well questions measuring the same underlying concept on the same occasion agree with each other. For example, two questions that measure different aspects of depression should agree with each other in that the same individual should score high or low on both. Agreement is measured quantitatively by correlations. For questionnaires with many questions measuring the same concept, a large number of correlations would be required to check the agreement of every question with every other question, and some summary of these correlations would be needed. Cronbach's alpha is a statistical measure which provides this. It is a single figure which summarizes the average correlation between all pairs of questions in a questionnaire. Cronbach's alphas should be above 0.7 and ideally between 0.8-0.9.[1] The reliability of each question can be assessed by calculating the alpha with and without it. Significant increases in alpha without the question would suggest that the question is not

---

[1]Above 0.9 suggests that the questionnaire may have too many questions and some could be eliminated (Streiner et al, 1995).

measuring the same thing as the other questions, and should be removed. Studying the effect of each question in this way is called Item Analysis.

Validity

Validity refers to the extent to which the measurement provided by an instrument agrees with the correct measurement. Instruments may be reliable but not valid, if they consistently give the same (but wrong) measurement and so both reliability and validity must be measured to assess instrument accuracy. There are two aspects of validity to be considered when testing a questionnaire:

Content validity

This refers to whether the instrument is considered by experts to be appropriate for measuring what it is supposed to measure. Part of content validity is whether experts believe that the questionnaire covers all the important aspects of the concept being studied. In the course of this study and the previous qualitative study we consulted two groups of 'experts.' The first group was mental health professional (including ICMC staff and JHU faculty) who helped us choose the instruments to adapt. The second group was the local population, through the qualitative study; the adapted instruments were chosen to match as closely as possible the psychosocial issues that emerged in that study.

Criterion validity

This refers to the agreement between the questionnaire and an external measure (criterion) of the same construct known to be accurate; In other words, comparing the questionnaire with a 'gold standard.' In this study we utilized an 'alternative gold standard' by relying on local community and religious leaders along with self-identification using the local terms and concepts for identifying significant psychosocial problems.