

# Interventions for Depression Symptoms Among Adolescent Survivors of War and Displacement in Northern Uganda

## A Randomized Controlled Trial

Paul Bolton, MBBS

Judith Bass, PhD

Theresa Betancourt, ScD

Liesbeth Speelman, MA

Grace Onyango, MA

Kathleen F. Clougherty, MSW

Richard Neugebauer, PhD

Laura Murray, PhD

Helen Verdelli, PhD

**T**HE WAR IN NORTHERN UGANDA remains one of the most violent and persistent complex humanitarian emergencies in the world. Over 1.8 million individuals, mainly ethnic Acholi, have been internally displaced during 20 years of conflict between the government of Uganda and the Lord's Resistance Army.<sup>1</sup> The Lord's Resistance Army has been accused of human rights abuses including mass violence, rape, and the abduction of more than 25 000 children.<sup>1</sup> Local populations have crowded into internally displaced persons camps where they face threats to their health and well being.<sup>1</sup> Prior research on children affected by armed conflicts documents increased risk of mental health problems ranging from adjustment difficulties to depression and anxiety disorders, including posttraumatic stress disorder.<sup>2-4</sup> Governmental and nongovernmental organizations (NGOs) intervene to address these problems, yet few interventions have been rigorously evaluated using a randomized controlled trial (RCT) design, either in northern Uganda or elsewhere.<sup>4</sup> The few RCTs in non-

**Context** Prior qualitative work with internally displaced persons in war-affected northern Uganda showed significant mental health and psychosocial problems.

**Objective** To assess effect of locally feasible interventions on depression, anxiety, and conduct problem symptoms among adolescent survivors of war and displacement in northern Uganda.

**Design, Setting, and Participants** A randomized controlled trial from May 2005 through December 2005 of 314 adolescents (aged 14-17 years) in 2 camps for internally displaced persons in northern Uganda.

**Interventions** Locally developed screening tools assessed the effectiveness of interventions in reducing symptoms of depression and anxiety, ameliorating conduct problems, and improving function among those who met study criteria and were randomly allocated (105, psychotherapy-based intervention [group interpersonal psychotherapy]; 105, activity-based intervention [creative play]; 104, wait-control group [individuals wait listed to receive treatment at study end]). Intervention groups met weekly for 16 weeks. Participants and controls were reassessed at end of study.

**Main Outcome Measures** Primary measure was a decrease in score (denoting improvement) on a depression symptom scale. Secondary measures were improvements in scores on anxiety, conduct problem symptoms, and function scales. Depression, anxiety, and conduct problems were assessed using the Acholi Psychosocial Assessment Instrument with a minimum score of 32 as the lower limit for clinically significant symptoms (maximum scale score, 105).

**Results** Difference in change in adjusted mean score for depression symptoms between group interpersonal psychotherapy and control groups was 9.79 points (95% confidence interval [CI], 1.66-17.93). Girls receiving group interpersonal psychotherapy showed substantial and significant improvement in depression symptoms compared with controls (12.61 points; 95% CI, 2.09-23.14). Improvement among boys was not statistically significant (5.72 points; 95% CI, -1.86 to 13.30). Creative play showed no effect on depression severity (-2.51 points; 95% CI, -11.42 to 6.39). There were no statistically different improvements in anxiety in either intervention group. Neither intervention improved conduct problem or function scores.

**Conclusions** Both interventions were locally feasible. Group interpersonal psychotherapy was effective for depression symptoms among adolescent girls affected by war and displacement. Other interventions should be investigated to assist adolescent boys in this population who have symptoms of depression.

**Trial Registration** clinicaltrials.gov Identifier: NCT00280319

JAMA. 2007;298(5):519-527

www.jama.com

Western societies that have included nonpharmaceutical treatments have generated mixed results.<sup>5-9</sup> Neither a community-based single-session testi-

mony intervention<sup>5</sup> nor a clinic-based counseling and problem-solving treatment<sup>7</sup> have so far proved superior to the control condition. In contrast, trials of

For editorial comment see p 567.

**Author Affiliations** are listed at the end of this article. **Corresponding Author:** Paul Bolton, MBBS, Center for Refugee and Disaster Response, Johns Hopkins

Bloomberg School of Public Health, 615 N Wolfe St, Room E8646, Baltimore, MD, 21205 (pbolton@jhsph.edu).

**Box. Symptoms of Locally Described Syndromes and Chores/Activities Used to Generate Local Depression Symptom and Functional Impairment Scales****Depression-Like Syndromes***Par*<sup>a</sup>

Has lots of thoughts, wants to be alone, is easily annoyed, holds head, drinks alcohol, thinks about suicide, doesn't greet people, sits alone, has lots of worries, does not think straight, cannot do anything to help self, does not trust, mutters to self, insults friends, is disobedient, is weak, cries continuously, loses concentration in school<sup>b</sup>

*Two Tam*<sup>a</sup>

Has lots of thoughts, constantly worries, experiences body pain, feels that brain isn't functioning, thinks of self as being of no use, thinks about suicide, talks about problems, sits alone, has headaches, feels sad, does not care about living or dying, thinks of bad things, doesn't feel like talking, is forgetful, is weak, cries continuously, loses interest in school<sup>b</sup>

*Kumu*<sup>a</sup>

Has loss of appetite, feels pain in the heart, sits with cheek in palm, cries when alone, does not sleep at night, talks about problems, lies down all the time, has lots of worries, has headaches, feels cold, is weak, does not feel like talking, is disobedient

**Anxiety-Like Syndromes***Ma Lwor*

Clings to elders, thinks of self as having no future, constantly runs, dislikes noise, has fast heart rate, fears being alone, has loss of appetite, wants to be alone, does not sleep at night, drinks alcohol, doesn't greet people, thinks people are chasing him/her

**Syndrome of Maladaptive Socially Unacceptable Behaviors***Kwo Maraco*

Fights, uses bad language, is disrespectful, misbehaves, drinks alcohol, uninterested, deceitful, a rough person, uses drugs, disobedient, loses interest in school<sup>b</sup>

**Play and Chore Activities by Function and Sex****Girls**

Fetching water, washing clothes, digging, washing utensils, playing ball, sweeping house, smearing floor with mud, cooking food, traditional dance

**Boys**

Fetching water, digging, playing games, sweeping compound, playing football

<sup>a</sup>Combined to create the composite depression scale. The final scale included 35 signs and symptoms.

<sup>b</sup>The 2 school-related items were excluded from the final scales because many of the students were not school-going.

a cognitive behavioral therapy,<sup>9</sup> a psychoeducation treatment program,<sup>6</sup> and a previous trial in southwestern Uganda (conducted by the authors of this article with an intervention also from this article),<sup>8</sup> all found significant improvements in the psychological symptoms studied. However, these trials focused on adults. RCTs of mental health interventions to assist children and adolescents affected by war are also needed.

This article describes an RCT in 2 camps for internally displaced persons in northern Uganda. We investigated whether a therapy-based intervention (interpersonal psychotherapy for groups, [IPT-G]) and an activity-based intervention (creative play, [CP]) were effective for relieving mental health and psychosocial problems resulting from war and displacement among adolescents.

**METHODS****Site and Population**

Study participants were Acholi adolescents aged 14 to 17 years living in 2 camps

for internally displaced persons near Gulu town in northern Uganda. Each camp has a population exceeding 20 000. Economic activity and organized social and cultural life are minimal, although children have some access to schools.

**Symptom Assessment**

We developed local symptom assessment measures based on a prior qualitative study in both camps (P.B., T.B., and L.S., unpublished data, 2004). In that study, we identified 7 important mental health and psychosocial problems affecting children and adolescents from the viewpoint of the youths themselves, their caregivers (mothers or other adult relatives), and other knowledgeable local individuals. Among these problems, the research team and NGO staff selected 5 problems judged amenable to interventions within the NGOs' capacity. Three of these 5 problems—*two tam*, *kumu*, and *par*—are local depression-like syndromes and were selected because of a

previous randomized trial we conducted, which found such syndromes to be amenable to treatment provided by one of our collaborating partners (World Vision).<sup>8</sup> The fourth local syndrome is *ma lwor*, which appears to be an anxiety-like syndrome and *kwo maraco*, the fifth local syndrome, is a combination of maladaptive or socially unacceptable behaviors (BOX). The remaining 2 problems, one referring to a cluster of posttraumatic stress reactions with psychotic features and the other to a legitimate fear of future Lord's Resistance Army attacks, were not investigated further because our NGO partners and the research team felt that these problems were unlikely to be affected by the NGOs' current or planned interventions, including the interventions tested in this study.

Problems of alcohol abuse and sexual violence were widely reported—symptoms related to the use of alcohol are present in 3 of the 5 locally described syndromes and sexual aggression is rep-

resented under the locally described conduct syndrome of *ma lwor*. However, our purpose was to study the effects of interventions on locally defined syndromes, and distinct syndromes referring to these problems were not identified in the qualitative study. Note, however, that while symptoms of psychosis, posttraumatic stress disorder, drug and alcohol use, and trauma were not among the foci of the assessment, individuals who had these symptoms were not excluded from the study.

As in a previous study<sup>10</sup> we explored whether any existing instruments matched the local descriptions of these 5 selected problems. Because no extant instruments proved suitable, we created a local instrument consisting of questions on all the symptoms of all 5 local problems: the Acholi Psychosocial Assessment Instrument (APAI). APAI respondents were asked about the frequency of each symptom occurrence during the previous week. Responses were coded on a 4-point Likert scale ranging from 0 (never) to 3 (constantly). Questions were added regarding symptom duration, sociodemographics, duration of camp residence, and whether the adolescent was ever abducted.

Given the predominance of depression-like syndromes in the qualitative data and the effectiveness of IPT-G for treating depression symptoms elsewhere in Uganda,<sup>8</sup> introducing and testing IPT-G in this population seemed appropriate. Both the research team and our NGO partners also believed that an existing intervention, CP, was potentially effective for these same problems (and for the conduct syndrome of *kwo maraco*). Therefore, we designated the depression problems as the primary study outcomes: *two tam*, *kumu*, and *par* (Box). Because the anxiety (*ma lwor*) and conduct (*kwo maraco*) syndromes were also thought amenable to both IPT-G and CP, measures of these problems were retained as secondary outcomes.

#### **Instrument Development, Reliability and Validity**

The 3 local depression problems contain varying (but incomplete) combina-

tions of the cardinal *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition) symptoms of depression and related symptoms.<sup>11</sup> Accordingly, we developed a local depression symptom scale based on the summed scores on all items comprising the 3 syndromes (35 individual signs and symptoms). Scale reliability and validity were evaluated for a subsample (n = 178) of the adolescents interviewed for trial eligibility (N=667). Cronbach  $\alpha$  (a measure of internal reliability) was 0.92. Mental health professionals familiar with the local culture were not available to make clinical diagnoses. Therefore, we used an alternative method of exploring criterion validity.<sup>10</sup> We identified individuals among this group for whom youth-caregiver pairs concurred that the youth had at least 2 of the 3 local depression-like syndromes (cases) and pairs concurring that the youth had none of these syndromes (noncases). The APAI depression symptom scale (mean [SD]) was substantially and significantly higher for the cases (45.3, [13.6]) than for the noncases (15.6, [11.2]), thereby supporting concurrent validity. The trial eligibility threshold was derived from these scores: 32, which is approximately 1 SD below the case mean. A convenience subsample of 30 individuals were readministered the APAI 3 to 5 days later. Test-retest reliability for the depression symptom scale was 0.84.

#### **Clinical Significance of Scale Scores**

The trial eligibility symptom threshold was derived from the comparison of mean depression symptom scale scores for cases and noncases, operationalized in the manner described in the previous section. The threshold of 32 (1 SD less than the case mean) was selected as a reasonable lower bound for clinically significant presence of symptoms (maximum scale score, 105). To evaluate the clinical significance of this threshold from a Western clinical perspective, the scale scores were compared with results for the Strengths and Difficulties Questionnaire,<sup>12</sup> administered together with the APAI. The Strengths and Difficulties Questionnaire, used previ-

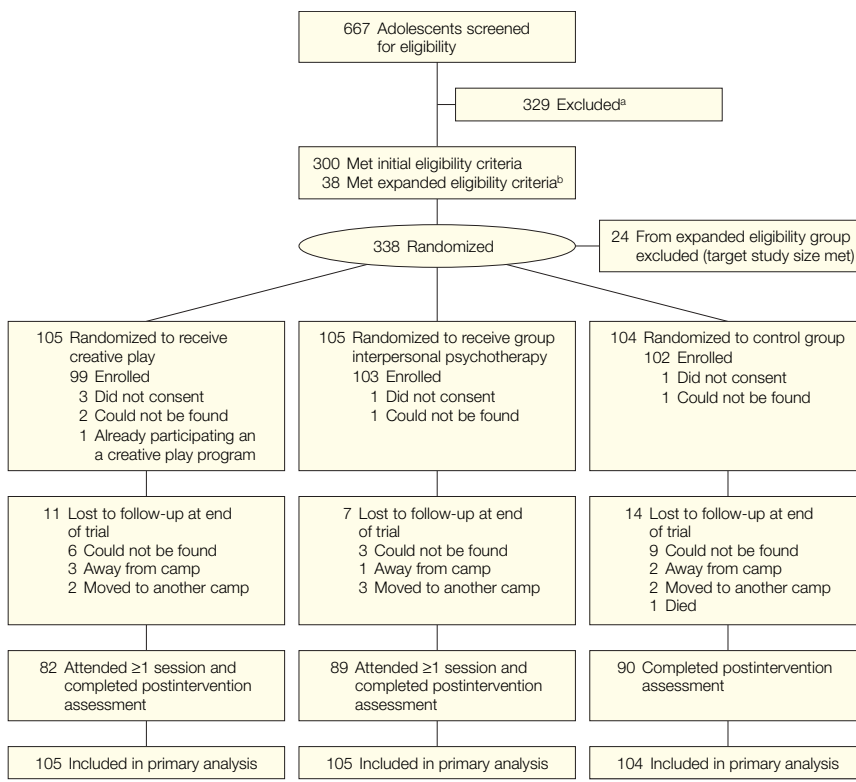
ously in a variety of cultures, is a general screening instrument for child emotional and behavioral problems.<sup>13-15</sup> It provides a determination of probable 'caseness' (based on scores)<sup>16</sup> but is not designed as a measure of treatment outcomes.<sup>17</sup> We compared the APAI scores for those individuals with positive screening results for the trial with their scores on the Strengths and Difficulties Questionnaire emotional problems subscale (for which scores of 7-10 points indicate probable "caseness")<sup>16</sup> to test whether the APAI and the cutoff criteria were identifying adolescents with significant mental health problems according to a standard measure (the Strengths and Difficulties Questionnaire).

#### **Functional Assessment**

We also developed a local function measure based on the qualitative data, using an approach similar to that used elsewhere.<sup>18</sup> The measure has 2 gender-specific scales, each reflecting common activities judged important for children and adolescents by youths and adult camp residents (Box). Impairment was rated on a Likert scale ranging from 0 (no more difficulty than most other boys/girls of the same age) to 4 (frequently unable). An overall function score comprised the sum of the individual item scores, and ranged from 0 to 36 for girls (9 activities) and 0 to 20 for boys (5 activities).

#### **Study Eligibility, Screening, Consent, and Randomization**

Screening and baseline assessments were conducted in the summer of 2005. Both interventions lasted from August 2005 through December 2005. Assessments were conducted pre- and post-intervention and change calculated for the local depression symptom scale. Sample size calculations were based on an expected 10-point difference in this change between either of the treatment groups and the controls. One hundred adolescents per group provided greater than 80% power ( $P=.05$ ), assuming SDs of less than 25 points. Direct comparisons between intervention groups were inappropriate given

**Figure.** Flowchart of Study Participants

<sup>a</sup>Did not meet inclusion criteria (described in Study Eligibility, Screening, Consent, and Randomization).

<sup>b</sup>Expanded eligibility criteria included reducing the cutoff score on the depression-severity screener.

differences in the size and sex ratio that made up the intervention groups. Consequently, the study was not powered for between-intervention comparisons.

Trial eligibility was based on a 2-stage screening process. In stage 1, 10 interviewer supervisors contacted local leaders, teachers, community workers, and adolescents to create a list of youths believed to have at least 1 of the local depression-like problems. In stage 2, the instrument was administered to these youths by 20 trained interviewers. Supervisors observed 10% of interviews, reviewed all completed instruments, and visited all adolescents not interviewed to confirm unavailability or refusal. Interviewers were used (rather than self-completion of forms) because of concerns about the rate of literacy among the respondents, given frequent interruptions in schooling due to the war.

Trial-eligible interviewees were aged 14 to 17 years, scored greater than 32

on the depression symptom scale and greater than 0 on the function scale, had symptoms for at least 1 month, and resided in the camps during the preceding month. Exclusion criteria were inability to be interviewed due to a cognitive or physical disability, or severe suicidal ideation or behavior.

Informed consent was initially obtained only for the screening interview. Eligible youths were then randomly assigned to a study group. Random allocation was done by computerized generation of a random number between 1 and 400 for each eligible participant, ordering them by number and assigning the first third to IPT-G, the second third to CP and the final third to the wait-control group. Each eligible youth was then visited by a facilitator (if allocated to IPT-G or CP) or other NGO staff (if allocated to controls) and a second informed consent process was done, this time for participation in the random-

ized trial. On this occasion informed consent included advising each youth of the study group to which he or she had been allocated. Our NGO partners had previously agreed to provide/continue on a permanent basis whichever intervention proved effective. Individuals assigned to the wait-control group were told they would be first to receive whichever intervention (if any) proved effective.

Of the total sample screened (N=667), 300 individuals met original inclusion criteria, were stratified by camp and sex, and randomized to a study group. Of these 300, 290 were enrolled in the study. Of the remaining 10 individuals, 1 was already involved in the CP program in a neighboring camp, 4 could not be located, and 5 refused. To meet our original sample size (300), we randomized an additional 38 individuals whose depression symptom scores were between 28 and 31 points. This relaxation of a trial eligibility criterion is acceptable when study design consequences are minimal.<sup>19</sup> The first 14 individuals all consented and therefore, the remainder were not approached. The participation rate was 96.8% (304 enrolled of 314 recruited; FIGURE).

### Postintervention Assessment

The study instrument was re-administered to 282 (90%) of the original 314 participants within 1 month of completing both interventions. Interviewers were blinded to interviewees' intervention status.

### Interventions

Both interventions comprised 16 weekly group meetings, lasting 1.5 to 2 hours each. These were preceded by 1 or 2 individual meetings in which the intervention was explained and (in the case of the IPT-G intervention) a treatment plan was generated by the IPT-G facilitators. Interpersonal psychotherapy, a time-limited, manualized intervention, was developed in the United States for ambulatory, nonpsychotic unipolar depression<sup>20</sup> and later adapted for use with adolescents with depression, individually<sup>21</sup> and in groups.<sup>22</sup> Interpersonal psychotherapy assumes that depressive epi-



sodes are triggered by difficulties in 1 or more of 4 interpersonal areas: grief, interpersonal disputes, role transitions, and interpersonal deficits. Interpersonal psychotherapy focuses on improving depressive symptoms and functioning by identifying the interpersonal problem(s) most relevant to the current depression and then assisting the individual in building skills to manage those problems. Treatment using IPT-G was first facilitated in Africa among adults with depression in southwest Uganda.<sup>8</sup> Based on this experience and our preliminary qualitative work in northern Uganda (P.B., T.B., and L.S., unpublished data, 2004 [available from authors at request]), IPT-G's focus on interpersonal triggers and group relationship building seemed compatible with Acholi culture.

The IPT-G intervention was implemented by World Vision Uganda. There were 12 IPT-G groups, each led by a facilitator with 2 weeks of on-site training by Columbia University faculty (K.F.C.). A treatment manual specifying IPT-G strategies and techniques was adapted for local use (K.F.C., H.V., and Myrna Weissman, PhD, Columbia University, New York State Psychiatric Institute, unpublished data, 2006). Groups consisted of 6 to 8 adolescent boys or girls and a facilitator of the same sex. Facilitators received weekly direct supervision by World Vision Uganda staff with prior IPT-G experience. These supervisors in turn received weekly phone supervision from the US-based trainer and were in weekly contact with study personnel. IPT-G supervisors also provided weekly written reports that were reviewed and discussed with study staff during the phone meetings for adherence to the treatment model and to monitor human subjects protection.

Use of CP refers to the creative play format developed by War Child Holland for war-affected youth. Its premise is that a youth's resilience will be strengthened by verbal and nonverbal expression of thoughts and feelings through age-appropriate creative activities such as, songs, art, role plays, music, sports, games and debates.<sup>23</sup> Each activity serves specific psychosocial goals and activities are selected from the War Child Holland

manual accordingly. For example, if the goal is to build trust with peers, the activities may require the adolescents to work collaboratively. After the activities, facilitators lead discussions on what the participants and facilitators thought about the activity as a means of drawing real-life lessons. CP was provided to 4 groups (2 groups/camp; both sexes represented in each group; 25-30 adolescents in each group), consistent with its routine implementation. The same 2 individuals (1 man, 1 woman) from War Child Holland jointly facilitated all CP sessions for all groups. Supervision was provided weekly or bi-monthly by the War Child Holland psychosocial specialist who regularly shared supervision reports with the US-based study personnel. These reports were reviewed and discussed with study staff during bimonthly phone meetings to ensure adherence to the treatment model and to monitor human subjects protection.

When implementing CP in different countries, some War Child Holland programs have added to the basic CP model by including various therapy-based elements in the postactivity discussion, according to local capacity. War Child Holland staff in northern Uganda had originally planned to do this as well. However, for this study, it was decided to use only the standard activity-based CP model in which postactivity group discussions focus on building skills. In this way, the study maximized the contrast between a psychotherapy-based intervention (IPT-G) and one focusing on group activities (CP). Participants could discuss their problems with the CP facilitators after the group meetings and this occurred on a few occasions for distressed participants (a young woman whose baby had died, and a boy who had lived with rebels in the bush), but these discussions were not a routine part of the intervention. Otherwise, the nature of the standard CP intervention was the same as that typically used by War Child Holland in other countries with 2 exceptions: (1) inclusion of adolescents as old as aged 17 years (usually limited to 15 years); and; (2) limiting CP participation, normally open to all children, to adolescents with depression symptoms.

## Analysis

Baseline characteristics of the study groups were compared using  $\chi^2$  tests for categorical data and *t* tests for continuous data. The primary outcome measure of intervention effectiveness was within-subject change in the depression symptom scale scores between baseline and postintervention assessment. The secondary outcome measures were the corresponding scale scores for the local anxiety and conduct syndrome symptoms and the function measure. We compared separately the mean within-subject change in the symptom and function scores between the wait-list control group and the individual interventions.

Stata statistical software version 9.0 (STATA Corp, College Station, Texas) was used for descriptive analyses and SAS version 9.1 (SAS Institute Inc, Cary, North Carolina) for grouped analyses. We used a random effects model to estimate intervention effect on change in scale scores, adjusting for the clustering within intervention groups by controlling for within-group correlated observations and between group variability.<sup>24</sup> Regression coefficients and *P* values were computed for continuous measures. Statistical significance was set at *P* < .05, 2-tailed and expressed as a 95% confidence interval (CI).

Analyses are presented for the intent-to-treat samples (*n* = 314), which includes all trial-eligible individuals randomized to one of the trial conditions and approached for inclusion, whether or not they consented. For individuals who were not reinterviewed, baseline scores were used as their postintervention scores, thereby imputing no change. Analyses were also conducted on a completer sample (*n* = 261), which included only individuals interviewed before and after the intervention and who (for the intervention groups) attended any sessions. This was done to determine if the results were substantially different for completers and as a check on whether our imputation method for noncompleters resulted in underestimates in sample variance.

The study was approved by the Boston University Institutional Review

**Table 1.** Study Sample Characteristics at Trial Baseline (n = 314)

Characteristic	Group Interpersonal Psychotherapy (n = 105)	Creative Play (n = 105)	Wait-List Controls (n = 104)	P Value <sup>a</sup>
No. (%) of participants				
Girls	60 (57)	61 (58)	59 (57)	.98
Currently enrolled in school	69 (66)	71 (68)	71 (68)	.88
History of abduction	42 (40)	48 (46)	41 (39)	.59
Mean (SD)				
Age, y	15.0 (1.1)	14.7 (1.0)	15.2 (1.2)	.01
Education, y	5.0 (1.5)	5.1 (1.4)	5.3 (1.3)	.47
Time in camp, y	5.0 (3.0)	5.5 (3.6)	5.0 (3.2)	.33
Local depression symptom, score <sup>b</sup>	43.5 (10.1)	44.2 (11.2)	44.2 (10.8)	.85
Function for girls, score <sup>c</sup>	11.8 (7.0)	11.3 (6.5)	10.7 (6.9)	.69
Function for boys, score <sup>c</sup>	6.8 (3.8)	7.1 (4.0)	8.2 (4.1)	.20

<sup>a</sup>Analysis of variance used to evaluate mean differences across groups for continuous variables and  $\chi^2$  used for categorical data.

<sup>b</sup>Depression symptom score generated by summing the 35 items referring to symptoms of the 3 local depression-like problems.

<sup>c</sup>Function scale for girls comprises 9 activities; function scale for boys comprises 5 activities. These include sex-specific play and chore activities.

Board and by the Ugandan National Council on Science and Technology in Kampala, Uganda.

## RESULTS

### Baseline Characteristics

Of 314 adolescents approached for enrollment, 172 (55%) were in Awer camp and 142 (45%) in Unyama. At baseline, 180 (57%) were girls, 211 (67%) were enrolled in school, and 131 (42%) reported a history of Lord's Resistance Army abduction. Mean duration of camp residence was 5.2 years. Except for a slightly older age among wait-list controls, the 3 study groups did not vary significantly (TABLE 1). The average Strengths and Difficulties Questionnaire emotional problems subscale score was 8.1 points without difference by sex and group assignment and 256 (81.5%) scored 7 or more points, the cutoff for probable 'case-ness'. Results are presented unadjusted for age as adjustment resulted in no appreciable difference.

### Attendance

Of the 105 adolescents assigned to IPT-G, 91 attended at least 2 sessions (mean of 14.1 sessions; SD, 2.6) of whom 89 (86% of the total IPT-G sample) were assessed postintervention. Of the 105 youth assigned to CP, 84 attended at least 3 sessions (mean of 10.4 sessions; SD, 2.9) of whom 82 (83% of the total CP sample) were reassessed. Ninety individuals

among the wait control group were interviewed postintervention (87% of the total 104 wait-list control sample). Fourteen adolescents (13%) assigned to IPT-G and 21 (20%) assigned to CP did not attend any sessions. Both sets of nonattendees had the same mean depression scores and ages as the attendees but were more likely to be girls, to not attend school, and to have a history of abduction.

### Changes in Symptom Severity

All 3 study groups demonstrated a decline in mean depression symptom scores between baseline and postintervention (TABLE 2). The decline in the depression symptom scale among the IPT-G group was substantially and significantly greater than the wait-list control group. The adjusted difference in mean within-subject decline between the IPT-G and wait-list control groups was 9.79 points (95% CI, 1.66-17.93) for the entire intent-to-treat sample and 10.71 points (95% CI, 4.02-17.40) for the completers only (Table 2). The CP group evidenced a small decline in the depression symptoms scale, not differing significantly from that of wait-list controls. The intraclass correlation coefficient for depression scale scores was 0.099. Variances of the outcome measure among the intent-to-treat sample differed only slightly from that among the Completer sample (Table 2), indicating that the imputation scheme did not materially influence the risk of type I error.

We also assessed change in the local anxiety problem (*ma lwor*) and conduct problems (*kwo maraco*) symptom scores as secondary outcomes. The adjusted difference in mean within-subject decline between the IPT-G and wait-list control groups for *ma lwor* was small but significant: 2.16 points (95% CI, 0.84-3.48) while that for *kwo maraco* was not: 0.74 points (95% CI, -0.15-1.63). The CP group showed a slight nonsignificant increase (0.2 points) and decrease (0.4 points) for *ma lwor* and *kwo maraco*, respectively.

Exploratory analyses examined whether the effectiveness of either intervention varied by sex, age, abduction history, and duration of camp residence. In stratified analyses, among girls, IPT-G was greatly superior to the wait-list control condition in reducing symptom severity, while the results for boys were not statistically significant (TABLE 3). This effect also held for the completer sample, where girls' mean scores were 16.6 points lower (95% CI, 9.93-22.13) for those who received IPT-G compared with controls and IPT-G boys had scores 5.01 points lower (95% CI, -1.89-11.90) compared with controls. No significant differences emerged for the other variables or in any of the CP analyses.

### Recovery and Remission

Recovery was defined as a reduction of 50% or more of an individual's baseline

**Table 2.** Effectiveness of Interventions Compared With Controls in Reducing Depression Symptom Scores Among Intent-to-Treat and Completer Samples

Score Type	Intent-to-Treat				
	Control (n = 104)	Group Interpersonal Psychotherapy (n = 105)	P Value	Creative Play (n = 105)	P Value
Baseline depression symptom score, mean (SD) <sup>a</sup>	44.2 (10.8)	43.5 (10.1)	.62	44.2 (11.2)	.99
Follow-up depression symptom score, mean (SD)	37.3 (15.9)	27.8 (17.2)		40.6 (15.7)	
Difference from baseline to follow-up, mean (SD)	6.9 (16.0)	15.7 (17.0)		3.6 (18.2)	
Difference in adjusted mean score change (SE) [95% confidence interval] <sup>b</sup>		9.79 (4.15) [1.66 to 17.93]	.02	-2.51 [-11.42 to 6.39]	.58

Score Type	Completers only				
	(n = 90)	(n = 89)		(n = 82)	
Baseline depression symptom score, mean (SD) <sup>a</sup>	44.1 (10.6)	42.9 (10.3)	.45	44.1 (11.4)	.99
Follow-up depression symptom score, mean (SD)	36.1 (16.1)	24.9 (16.5)		39.6 (15.7)	
Difference from baseline to follow-up, mean (SD)	8.0 (17.0)	18.0 (16.5)		4.5 (19.0)	
Difference in adjusted mean score change (SE) [95% confidence interval] <sup>b</sup>		10.71 (3.4) [4.02 to 17.40]	.002	-3.90 (3.7) [-11.17 to 3.37]	.29

<sup>a</sup>P value refers to significance of difference in baseline scores between intervention and control groups.

<sup>b</sup>Estimates of difference between intervention change score and control change score and associated confidence intervals and p values calculated using random effects model adjusting for clustering and baseline depression scale scores. Intraclass correlation coefficient calculated at 0.099.

symptom severity score and remission as being at or less than a predefined cut-off score of 15.6 points (the mean score of the noncases in the validation study). Among wait-list controls, 14 individuals (13.5%) met criteria for recovery and 9 of total (8.9) met criteria for remission. The corresponding numbers for the CP and IPT-G groups were 13 (12.4%) and 39 (37.1%) for recovery and 7 of total (6.7%) and 30 of total (29.1%) for remission, respectively. Among the IPT-G sample, 25 (42%) of the girls and 14 (31%) of the boys met recovery criteria while 20 (33.9%) of the girls and 10 (22.7%) of the boys met remission criteria. Among 'completers', 14.6% and 7.3% of the CP participants and 42.7% and 34.5% of the IPT-G participants respectively, met the recovery and remission criteria.

#### Attendance and Group Leader Effects

The number of IPT-G sessions attended was associated with greater symptom reduction ( $\beta=1.3$ ;  $P=.001$ ). No association was found between number of CP sessions attended and symptom change. IPT-G effectiveness did not vary significantly by group facilitator but the limited number of groups (12) constrained analytic power. Facilitator effects could not be examined for CP because both facilitators worked with all 4 groups.

**Table 3.** Interventions Compared With Control in Reducing Depression Symptom Scores Among Intent-to-Treat Samples

	Difference in Scores (95% Confidence Interval) <sup>a</sup>	
	Group Interpersonal Psychotherapy vs Wait-List Controls	Creative Play vs Wait-List Controls
Girls	12.61 (2.09 to 23.14)	-1.61 (-12.51 to 9.29)
Boys	5.72 (-1.86 to 13.30)	-3.40 (-11.17 to 4.36)

<sup>a</sup>Data presented as estimate of difference (95% confidence interval), calculated using random effects model adjusting for clustering and baseline scores.

#### Change in Function

Mean within-subject function scores declined (ie, function improved) similarly across all 3 study groups among girls and boys. Among the individuals in the control group, the girls experienced a mean 1.9-point reduction (95% CI, -0.11 to 3.97) in their functional impairment scores and the boys experienced a 0.8-point mean reduction (95% CI, -0.39 to 2.07). This was similar to the amounts experienced by the CP participants with a mean 1.3-point reduction for girls (95% CI, -0.73 to 3.23) and 0.5-point reduction for boys (95% CI, -0.69 to 1.66). The IPT-G participants reported greater reductions but the overall differences compared with controls did not reach statistical significance. Among the IPT-G girls, the mean reduction was 3.2 points (95% CI, 1.51-4.87) while the boys' mean score was 1.3 points (95% CI, -0.02 to 2.93).

#### Effect of Using the Dunnett Adjustment

Proschan and Fullman have recommended using the Dunnett adjustment for statistical analysis of multiple comparisons for the same reasons that the bonferroni adjustment is used.<sup>25</sup> While that article refers to more comparisons than we conducted in this study, we nevertheless tested the effect of including the adjustment. We found no differences in the data and the adjustment factor itself was not significant.

#### COMMENT

This study found IPT-G (a psychotherapy-based intervention) superior to a wait-list control condition in reducing depressive symptoms among this adolescent population. Statistically significant improvement was limited to the girl participants in this study. CP was not superior to the wait-list control con-

dition. Neither IPT-G nor CP was effective in improving anxiety, conduct problems, or functioning among boys or girls, according to our measures.

Girls derived much greater improvement in depression symptoms from IPT-G than did boys. The amount of change among boys was not statistically significant when compared with controls, however the study was not originally powered to detect gender-specific differences. Previous studies of IPT-G among adolescents in Western countries have included too few boys to know whether this finding reflects the situation elsewhere.<sup>22</sup> It may be that boys are less willing to talk about emotional problems, particularly in a group format. The different comorbidity profiles among the boys and girls (boys had more substance use and posttrauma symptoms) may also have affected the effect of IPT-G, which has been shown to be less effective in individuals with depression and comorbid anxiety disorders.<sup>26</sup>

The lack of improvement among the CP group may be because it was originally designed for younger ages ( $\leq 15$  years) than our study population, who may be more like adults than children in this culture and context. This may also explain the effectiveness of an adult intervention like IPT-G. The absence of effectiveness among the CP group was unexpected, particularly since the CP facilitators received many positive comments from participants, caretakers, and teachers. Perhaps the participants wanted to please the facilitators, or were referring to other improvements of which the nature was not captured with our measures, or they were referring to the same natural improvement that occurred among the control group (but which they may have attributed to the intervention). Whatever the reason, we note that an activity-based intervention did not produce significant improvement in the outcomes assessed in the present study.

The lack of improvement in function for both interventions is in agreement with reviews that note that reduction in depressive symptoms does not necessarily translate into im-

proved functioning.<sup>27,28</sup> Another possibility is that functional improvement occurred sometime later, a result that characterizes some interpersonal psychotherapy trials in industrialized countries,<sup>29</sup> or that the lack of improvement reflected a limitation in our measures (see Study Limitations section).

The IPT-G group also showed a small but significant improvement in anxiety symptoms compared with controls, which is consistent with the improvement in depression symptoms and the close relationship between anxiety and depression. There was no such improvement among the CP group and conduct symptoms showed no improvement in either group compared with controls. Given that validity testing of the APAI focused on the depression syndromes, and that participants were screened into the study on the basis of depression symptoms only, we consider these results to be suggestive only. Further investigation is required.

The consent process for the randomized trial differed from standard RCT procedures, where full consent precedes random allocation. Our approach was based on prior experience among less-educated populations in Sub-Saharan Africa, where the concept of random assignment is not well understood and can result in withdrawal from the study by those allocated to the control group. Instead, we randomized potential participants after the screening and then informed them of their allocation as part of the trial-consent process. In this way, we expected to reduce both refusals and drop-outs—as long as refusals and drop-outs were minimal, selection bias would not be significant. Only 10 individuals randomized in this way did not participate. Of these, 5 were refusals. All 10 were included in the intention-to-treat analyses by assuming no change between pre- and postassessments.

### Study Limitations

While the lack of improvement among the CP groups might suggest that the facilitator and group effects were not important, we cannot conclude this because the same facilitators ran both

groups and because those groups were larger and of mixed gender, both of which would reduce these effects. As a result, we cannot speculate by comparison with CP what it was about IPT-G that was effective; whether it was the psychotherapy, the effect of meeting in small groups, or the nonspecific attention of the facilitator.

The lack of improvement in function among any of the study groups may be due, at least in part, to limitations in our measures. The items were taken solely from the qualitative data and limited to activities that were possible given the environment (a camp for internally displaced persons). This resulted in function scales containing only a few items (9 for girls and 5 for boys) and emphasizing activities and tasks and not capturing other types of interpersonal or social functioning targeted by the IPT-G and CP interventions.

In our prestudy power calculations, we did not take clustering into account, possibly resulting in some underestimation of the required sample size. However, the intraclass correlation coefficient analysis suggests minimal variance due to the clustering of individuals within intervention groups.

To our knowledge, this is the first RCT of psychosocial or mental health interventions among African adolescents affected by war, and one of only a few RCTs of psychological treatments for depression symptoms conducted in a developing country. In addition to our previous trial of IPT-G in southwestern Uganda,<sup>8</sup> are RCTs by Patel et al in India,<sup>7</sup> Araya et al in Chile,<sup>6</sup> and Igreja et al<sup>5</sup> in Mozambique; all focusing on adults. Two of these studies used individual-based psychotherapy treatments<sup>5,7</sup> and reported no effectiveness, whereas the 2 group-based studies found active treatment superior to controls. The current study also supports the potential of group-based therapies.

Despite the differences between the current study population and that of our previous study<sup>8</sup> of IPT-G (different culture, language, and ethnicity; adults vs adolescents; war-affected and displaced vs nondisplaced and at peace) there were



similarities in the results. IPT-G proved feasible in both sites in that acceptance was high, as was attendance throughout the intervention period. In both sites there was significant overall improvement in depression symptoms, compared with controls, with females improving substantially more than males. The extent of overall symptom improvement was less in this study than among the adults and function did not significantly improve as it did among the adults. While this may be due to the study limitations as described previously, it may also suggest that IPT-G is more suited for mature participants and more suited to girls across the age groups.

The failure of both IPT-G and CP to significantly assist boys in this study raises the question of whether other interventions may be needed to assist war-affected boys with depression symptoms. Since both group psychotherapy and activity-based interventions were not effective, some form of individual psychotherapy or an entirely different type of intervention may be indicated as the basis for a future trial.

This study suggests that effective psychological interventions can be feasible in poor, rural, and illiterate communities (even those affected by war) as are formal trials of such interventions. Where humanitarian and treatment programs screen on the basis of symptom severity, normal fluctuations in severity over time, and regression to the mean will ensure apparent improvement regardless of true intervention effectiveness. Such programs, therefore, require controlled study designs to accurately determine effectiveness.

**Author Affiliations:** Applied Mental Health Research Group, Center for International Health and Development, Boston University School of Public Health, Boston, Massachusetts (Drs Bolton, Bass, and Murray); Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Dr Bass); Department of Population and International Health, François-Xavier Bagnoud Center for Health and Human Rights, Harvard University School of Public Health, Boston, Massachusetts (Dr Betancourt); War Child Holland, Gulu, Uganda, (Ms Speelman); World Vision Uganda, Kampala, Uganda (Ms Onyango); New York State Psychiatric Institute, Columbia University, New York, New York (Ms Clougherty and Dr Verdelli); Division of Epidemiology, New York State Psychiatric Insti-

tute, Columbia University College of Physicians and Surgeons, and Gertrude H. Sergievsky Center, Faculty of Medicine, College of Physicians and Surgeons, Columbia University, New York, New York (Dr Neugebauer); Teachers College, Columbia University (Dr Verdelli); Dr Bolton is now with the Center for Refugee and Disaster Response, and Dr Bass is now with the Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

**Author Contributions:** Dr Bolton had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. **Study concept and design:** Bolton, Betancourt, Bass. **Acquisition of data:** Bolton, Bass, Betancourt. **Analysis and interpretation of data:** Bolton, Bass, Betancourt, Clougherty, Neugebauer, Murray, Verdelli. **Drafting of the manuscript:** Bolton, Bass, Betancourt, Speelman, Onyango, Clougherty, Neugebauer, Murray, Verdelli.

**Statistical analysis:** Bass, Betancourt, Neugebauer. **Study supervision:** Bolton, Betancourt, Speelman, Onyango, Clougherty, Murray, Verdelli. **Financial Disclosures:** None reported.

**Funding/Support:** This project was solely funded by World Vision and War Child Holland. Dr Neugebauer's contributions were funded by the Ruth and David Levine Foundation.

**Role of Sponsors:** Both organizations also provided material support including local staff time to conduct and supervise the interventions, assist in the logistical aspects of both the study design and interventions, and assisted in supervision of the interviewers. Neither organization was involved in the management or analysis of the data but did offer suggestions as to the interpretation of the results. These staff also reviewed and approved the manuscript.

**Additional Contributions:** We wish to thank those staff, local leaders, and the communities who welcomed us to work with them. Further, we wish to thank both organizations for their commitment to improving the impact of their humanitarian efforts. Both organizations have already begun to use the results of this study to inform their current programs in northern Uganda. While not funding the study directly, the methodology described in this article was developed with the support of the US Agency for International Development's Torture Victims Fund.

## REFERENCES

- Coalition to Stop the Use of Child Soldiers. Child Soldiers Global Report 2004. [http://www.reliefweb.int/rw/lib.nsf/db9005ID/LHON-66UJKU/\\$FILE/child\\_soldiers\\_CSC\\_nov\\_2004.pdf?OpenElement](http://www.reliefweb.int/rw/lib.nsf/db9005ID/LHON-66UJKU/$FILE/child_soldiers_CSC_nov_2004.pdf?OpenElement) Accessed July 9 2007.
- Barenbaum J, Ruchkin V, Schwab-Stone MS. The psychosocial aspects of children exposed to war. *J Child Psychol Psychiatry*. 2004;45(1):41-62.
- Lustig SL, Kia-Keating M, Grant Knight W, et al. Review of child and adolescent refugee mental health. *J Am Acad Child Adolesc Psychiatry*. 2004;43(1):24-36.
- Stichick T. The psychosocial impact of armed conflict on children. *Child Adolesc Psychiatr Clin N Am*. 2001;10(4):797-814.
- Igreja V, Kleijn WC, Schreuder B, Van Dijk JA, Verschuur M. Testimony method to ameliorate post-traumatic stress symptoms. *Br J Psychiatry*. 2004;184(3):251-257.
- Araya R, Rojas G, Fritsch R, Gaete J, Simon G, Peters TJ. Treating depression in primary care among low-income women in Santiago, Chile: a randomized, controlled trial. *Lancet*. 2003;361(9362):995-1000.
- Patel V, Chisholm D, Rabe-Hesketh S, Dias-Saxena F, Andrew G, Mann A. Efficacy and cost-effectiveness of a drug and psychological treatment for common mental disorders in general health care in Goa, India. *Lancet*. 2003;361(9351):33-39.

- Bolton P, Bass J, Neugebauer R, et al. Results of a clinical trial of a group intervention for depression in rural Uganda. *JAMA*. 2003;289(23):3117-3124.
- Sumathipala A, Hewege S, Hanwella R, Mann AH. Randomized controlled trial of cognitive behaviour therapy for repeated consultations for medically unexplained complaints: a feasibility study in Sri Lanka. *Psychol Med*. 2000;30(4):747-757.
- Bolton P. Cross-cultural validity and reliability testing of a standard psychiatric assessment instrument without a gold standard. *J Nerv Ment Dis*. 2001;189(4):238-242.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association; 2000.
- Goodman R. Strengths and difficulties questionnaire. *J Child Psychol Psychiatry*. 1997;38(5):581-586.
- Kashala E, Lundervold A, Sommerfelt K, Tylleskar T, Elgen I. *Co-Existing Symptoms and Risk Factors Among African School Children with Hyperactivity-Inattention Symptoms in Kinshasa* [published online]. Congo, Africa: Eur Child Adolesc Psychiatry; 2006.
- Samad L, Hollis C, Prince M, Goodman R. Child and adolescent psychopathology in a developing country. *Int J Methods Psychiatr Res*. 2005;14(3):158-166.
- Thabet AA, Tischler V, Vostanis P. Maltreatment and coping strategies among male adolescents living in the Gaza Strip. *Child Abuse Negl*. 2004;28(1):77-91.
- Goodman R, Renfrew D, Mullick M. Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *Eur Child Adolesc Psychiatry*. 2000;9(2):129-134.
- Goodman R, Scott S. Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful? *J Abnorm Child Psychol*. 1999;27(1):17-24.
- Bolton P, Tang AM. An alternative approach to cross-cultural function assessment. *Soc Psychiatry Psychiatr Epidemiol*. 2002;37(11):537-543.
- Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials*. 3rd ed. New York, NY: Springer; 1998:151.
- Weissman M, Markowitz J, Klerman G. *Comprehensive Guide to Interpersonal Psychotherapy*. New York, NY: Basic Books; 2000.
- Mufson L, Weissman MM, Moreau D, Garfinkel R. Efficacy of interpersonal psychotherapy for depressed adolescents. *Arch Gen Psychiatry*. 1999;56(6):573-579.
- Mufson L, Gallagher T, Dorta KP, Young JF. A group adaptation of interpersonal psychotherapy for depressed adolescents. *Am J Psychother*. 2004;58(2):220-237.
- Kalksma-van Lith B. State of the Art in Psychosocial Interventions with Children in War-affected Areas. War Child Holland 2005. <http://admin2.warchild.nl/uploadedfiles/979.pdf>. Accessed February 3, 2007.
- Omar RZ, Thompson SG. Analysis of a cluster randomized trial with binary outcome data using a multi-level model. *Stat Med*. 2000;19(19):2675-2688.
- Proschan MA, Follmann DA. Multiple comparisons with control in a single experiment versus separate experiments: why do we feel differently? *Am Stat*. 1995;49:144-149J.
- Young JF, Mufson L, Davies M. Impact of comorbid anxiety in an effectiveness study of interpersonal psychotherapy for depressed adolescents. *J Am Acad Child Adolesc Psychiatry*. 2006;45(8):904-912.
- Hirschfield RM, Montgomery SA, Keller MB, et al. Social functioning in depression: a review. *J Clin Psychiatry*. 2000;61(4):268-275.
- Lin EH, VonKorff M, Russo J, et al. Can depression treatment in primary care reduce disability? *Arch Fam Med*. 2000;9(10):1052-1058.
- Weissman MM, Klerman GL, Prusoff BA, Sholomskas D, Padian N. Depressed outpatients. *Arch Gen Psychiatry*. 1981;38(1):51-55.