

The Development and Psychometric Properties of the Humanitarian Emergency Settings Perceived Needs (HESPER) Scale

Maya Semrau, MSc, Mark van Ommeren, PhD, Monica Blagescu, MA, Andre Griekspoor, MD, MPH, Louise M. Howard, PhD, Mark Jordans, PhD, Heidi Lempp, PhD, Anita Marini, PhD, Jon Pedersen, MA, Isabelle Pilotte, MSc, Mike Slade, PhD, and Graham Thornicroft, PhD

Needs assessments in humanitarian settings (i.e., places in which a large part of the population is at risk of dying or experiencing immense suffering) are vital in enabling effective and efficient emergency relief. However, current needs assessments are often far from ideal; indeed, in 2009, heads of 26 large humanitarian donor agencies signed a letter to the United Nations asking for an improvement in the area of needs assessment (J. Isbister, G. Weinberger, J.-P. Loir, et al., unpublished letter, 2009).

There have also been repeated recommendations for increased participation of affected populations in humanitarian assessment.¹⁻⁶ People's participation in assessment is seen as a right and as essential for optimizing resource allocation, program design, and population empowerment.⁶ It increases the likelihood that interventions are based on needs as expressed by the affected population. The international humanitarian community's focus on participation is exemplified by the fact that the recently revised, influential *Sphere Handbook* (^{5,6}) on standards for humanitarian aid emphasizes the involvement of affected people.

Participation is recommended throughout the assessment, design, monitoring, and evaluation program cycle.^{1,3-5} Additionally, in a recent ranking exercise for research priorities in the area of mental health and psychosocial support, 3 of the 10 most highly prioritized research questions in humanitarian settings included the participation of affected populations; the identification of affected populations' stressors was ranked as top priority.⁷ Related to this is the notion of accountability within the international humanitarian response, including that humanitarian action should be accountable to affected populations.⁴

Within this framework of increased participation and accountability, it has been

Objectives. We developed the Humanitarian Emergency Settings Perceived Needs (HESPER) Scale, a valid and reliable scale to rapidly assess perceived needs of populations in humanitarian settings in low- and middle-income countries.

Methods. We generated items through a literature review; reduced the number of items on the basis of a survey with humanitarian experts; pilot-tested the scale in Gaza, Jordan, Sudan, and the United Kingdom; and field-tested it in Haiti, Jordan, and Nepal.

Results. During field-testing, intraclass correlation coefficients (absolute agreement) for the total number of unmet needs were 0.998 in Jordan, 0.986 in Haiti, and 0.995 in Nepal (interrater reliability), and 0.961 in Jordan and 0.773 in Nepal (test-retest reliability). Cohen's κ for the 26 individual HESPER items ranged between 0.66 and 1.0 (interrater reliability) and between 0.07 and 1.0 (test-retest reliability) across sites. Most HESPER items correlated as predicted with related questions of the World Health Organization Quality of Life-100 (WHOQOL-100), and participants found items comprehensive and relevant, suggesting criterion (concurrent) validity and content validity.

Conclusions. The HESPER Scale rapidly provides valid and reliable population-based data on perceived needs in humanitarian settings. (*Am J Public Health.* 2012;102:e55-e63. doi:10.2105/AJPH.2012.300720)

recommended that the assessment of perceived needs be used to inform project design, monitoring, and evaluation,^{1-5,8,9} and perceived needs are considered a key determinant of psychosocial well-being.^{1,8,10} Perceived needs are defined here as needs expressed by members of the affected population themselves. They are thus problem areas for which people would likely want help. In the humanitarian field, perceived needs are still assessed mostly through rapid participatory assessments in the early phase of a crisis; these assessments tend to involve gaining qualitative data from selected stakeholders through focus groups or key respondent interviews.¹¹ Although certainly valuable, such assessments cannot provide a population-level picture. Most population-based quantitative assessments are of "objective" indicators, such as mortality rates, malnutrition rates, or livelihood data.¹²⁻¹⁴

These indicators are often defined by outsiders (i.e., nonmembers of the affected population) and do not quantify the prevalence and distribution of needs as perceived by members of the population themselves.

With a few exceptions,¹⁵⁻¹⁷ assessment tools in the humanitarian field tend to have unknown psychometric properties (i.e., indices of validity and reliability). Without published psychometric properties, it is unknown to what extent assessment tools are fit for purpose.

To address these gaps, we developed a method and instrument to rapidly and quantitatively assess perceived needs in emergency-affected populations—the Humanitarian Emergency Settings Perceived Needs (HESPER) Scale.¹⁸ We describe the development and psychometric properties of the scale.

OVERVIEW OF HESPER SCALE

The HESPER Scale assesses the perceived physical, social, and psychological needs of the general adult population in humanitarian settings during conflict or other disasters in low- and middle-income countries. Perceived needs are assessed on the HESPER Scale across 26 need items, which each includes a short item heading, as well as an accompanying question. Examples of need items include “Place to Live In” (“Do you have a serious problem because you do not have an adequate place to live in?”) and “Education for Your Children” (“Do you have a serious problem because your children are not in school or are not getting a good enough education?”). Ratings are then made for each need item according to unmet need (or serious problem; “1” rating), no need (or no serious problem; “0” rating), or no answer (i.e., refused, not known, or not applicable; “9” rating). From among the items that participants have rated as unmet needs, they are asked to rank their 3 most serious problems (hereafter referred to as priority ratings). This may enable prioritization of needs and emergency relief to those areas where it is perceived to be needed most. Participants are also asked to name any additional unmet needs not already listed. A total score of unmet needs can be calculated by adding up the number of items rated as serious problems.

The HESPER Scale was modeled on a mental health instrument, the Camberwell Assessment of Need Short Appraisal Schedule (CANSAS),¹⁹ which has well-established reliability and validity.^{20,21} The CANSAS has been modified successfully for numerous populations^{22–25} and adapted for use in several countries.²⁶ It has been used on a wide range of populations, including asylum seekers and refugees in the United Kingdom^{27,28} and torture victims in centers of the International Rehabilitation Council for Torture Victims in several countries.

METHODS

We developed the HESPER Scale over 3 phases (Figure 1):

- Phase 1 (2008): development of a draft scale through a process of item generation and item reduction, based first on a literature

review and second on a survey with humanitarian experts.

- Phase 2 (2009): pilot-testing of the draft scale—in Jordan with displaced Iraqi people, in Gaza and Sudan with the general adult population, and in the United Kingdom with refugees from the Democratic Republic of the Congo—to assess the scale’s feasibility, intelligibility, and cultural applicability, and to establish the suitability of training materials.
- Phase 3 (2010): field-testing of the revised draft scale—in Jordan with displaced Iraqi people, in Haiti with people living in post-earthquake displacement camps, and in Nepal with Bhutanese refugees—to assess its psychometric properties (i.e., validity and reliability).

Procedure

A steering committee and advisory group composed of international experts guided the development of the HESPER Scale.

Phase 1—development of draft scale. We developed the first draft scale²⁹ through a process of item generation and item reduction. We generated an item pool of 38 items by extracting items from gray and peer-reviewed literature directly documenting emergency-affected people’s views of perceived needs, such as previous humanitarian needs assessments, existing assessment reports of non-governmental organizations, and published journal articles on perceived needs (B. Poudyal, T. Erni, A. Jonathan, et al., unpublished data, 2007; S. B. Thapa and E. Hauff, unpublished data, 2007).^{8,30–38} We included only items that were mentioned at least twice in any of these sources.

We then selected and reduced need items into the draft scale on the basis of a survey with a wide range of purposively sampled general and psychosocial humanitarian experts across the world (24 men and 19 women), as well as 6 national aid workers in Sierra Leone. The survey included both quantitative and qualitative responses; participants rated the need items that had been compiled during the item generation stage on an 11-point scale (0–10) of importance for inclusion into the scale, and suggested additional perceived need items that they considered important for inclusion. In addition, participants were encouraged to provide any further comments or feedback.²⁹ We

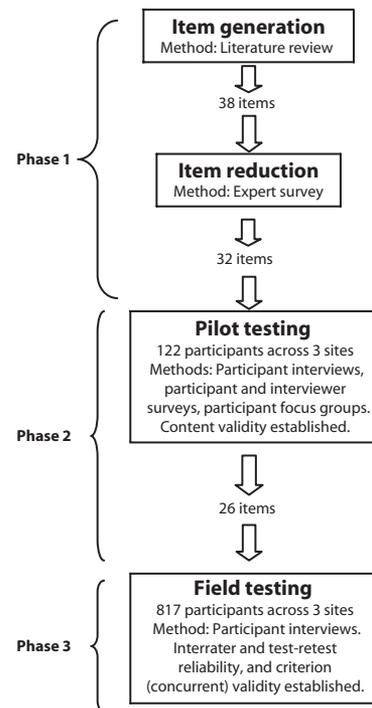


FIGURE 1—Development of the Humanitarian Emergency Settings Perceived Needs (HESPER) Scale over 3 phases.

drafted training materials to accompany the scale.

Phase 2—pilot-testing. We then pilot-tested the draft HESPER Scale in 3 relevant settings, after pretesting it in the United Kingdom with 7 refugees from Democratic Republic of the Congo who had been resettled from refugee camps in Zambia. Pilot-testing was a learning exercise to understand the scale’s feasibility, intelligibility, and cultural applicability,³⁹ as well as assessing methodologies for subsequent field-testing.

We employed convenience sampling to recruit participants in the 3 pilot sites, with interviewers identifying and selecting participants. The following were interviewed: 40 Iraqis displaced following the 2003 invasion of Iraq (interviewed in Amman, Jordan in June 2009), 40 members of the local population in Gaza City (October 2009), and 42 members of the local population in Juba, Sudan (December 2009). All participants were at least 18 years old.

Interviews were conducted in participants' homes in one-to-one assessments; between 4 and 7 local interviewers (of whom 53.3% were women and 46.7% men) conducted interviews in the local Arabic dialect at each of the 3 pilot sites. We previously trained interviewers for 1 to 1½ days in administering the HESPER Scale. Interviewers administered the draft HESPER Scale to participants, as well as a survey in which participants were asked about any missing items and the intelligibility of the draft scale. For a subsample (20 each in Jordan and Gaza and 18 in Sudan), a second interviewer acted as silent rater to assess interrater reliability. Interviewers also invited participants to take part in a focus group discussion, in which participants reported on the intelligibility, (cultural) acceptability, relevance, and comprehensiveness of the scale's items, as well as the suitability of the content and concepts. We conducted 4 focus groups (2 for men, 2 for women) in each of the 3 pilot sites; 15 participants chose to take part in Jordan, 33 in Gaza, and 12 in Sudan. Interviewers completed an interviewer survey, in which they provided feedback on the intelligibility of the HESPER Scale and training materials, and on whether they experienced any difficulties in conducting the interviews.

Phase 3—field-testing. We then field-tested the revised HESPER Scale with larger samples in 3 relevant humanitarian settings to assess its psychometric properties and to estimate the level of perceived needs in these settings (here we focus on the scale's psychometric properties only). In total, 269 Iraqi participants displaced following the 2003 invasion of Iraq were interviewed in Jordan (Amman, Zarqa, Irbid, and Madaba) in July 2010, 279 people living in displacement camps following the January 2010 earthquake were interviewed in Haiti (Champs de Mars and Bolosse camps in Port-au-Prince and Pinchinat camp in Jacmel) in September 2010, and 269 Bhutanese refugees were interviewed in Nepal (Beldangi-II camp in Jhapa district) in October and November 2010.

Project materials were translated by back-translation methods before field-testing commenced; a bilingual translator first translated materials into the local language, another translator then translated the materials back into English, and the 2 versions were compared

to identify and resolve any mistakes in the translation.⁴⁰

To determine sample sizes for field-testing of the psychometric properties of the scale, we performed a calculation for test–retest reliability on the basis of previous psychometric testing of the different CANSAS versions. This showed the required minimum sample size for test–retest reliability to be 69 per site to give power $(1 - \beta)$ of 0.8, using a P value of .05, a minimum acceptable level of test–retest reliability (intraclass correlation coefficient) of 0.6, and a predicted test–retest reliability (intraclass correlation coefficient) of 0.7.⁴¹ This sample size also allowed the detection of correlations for criterion (concurrent) validity of at least $r = 0.3$ with power $(1 - \beta)$ of 99%, or $r = 0.2$ with power $(1 - \beta)$ of 83%. Furthermore, we performed a calculation for interrater reliability on the basis of findings made during previous pilot-testing of the HESPER Scale. This showed the required minimum sample size for interrater reliability to be 39 per site in order to give power of 0.8, using a P value of .05, a minimum acceptable level of interrater reliability of 0.7, and a predicted interrater reliability of 0.8.⁴¹

We employed different sampling methods in the 3 sites according to what was appropriate and feasible. Iraqi participants in Jordan were recruited through a multistage cluster sampling design, involving 30 clusters of city districts. The sample was geographically representative of Iraqis living in Jordan, with around 75% of the sample in Amman (23 clusters) and around 25% (7 clusters) in other governorates (4 in Zarqa, 2 in Irbid, and 1 in Madaba). In Haiti, we purposively selected 3 displacement camps as study sites to fit in with the implementing agency's programs. Within camps, we selected participants by using a 2-stage systematic random sampling method, the first stage being households and the second stage being individuals within households. Both in Jordan and Haiti, we employed random-walk methods to recruit households within clusters or camps; we then randomly selected individuals within chosen households by using a random-number Kish Table.⁴² In Nepal, we employed simple random sampling methods to recruit participants; we obtained a list of randomly selected Bhutanese refugees living in Beldangi-II camp from the Office of

the United Nations High Commissioner for Refugees.

In each setting, between 6 and 12 local interviewers (57.7% were men and 42.3% were women) conducted interviews in one-to-one assessments in participants' homes (in Arabic in Jordan, Haitian Creole in Haiti, and Nepali in Nepal); the interviewers had previously been trained for 2 days (including a half-day pilot) in administering the HESPER Scale. Interviewers were recruited by the local collaborating organizations following an interview process, and were supervised by a local team leader. To measure the HESPER Scale's interrater reliability, a second interviewer acted as silent rater for 46 participants in Jordan, 44 in Haiti, and 42 in Nepal. To assess test–retest reliability of the scale, 70 and 73 participants in Jordan and Nepal, respectively, were interviewed a second time 1 week after the first interview by the same interviewer who had interviewed them before. We did not assess test–retest reliability in Haiti as it was considered too burdensome for local people in this intense humanitarian setting. We established criterion (concurrent) validity of the HESPER Scale by comparing 15 of its 26 individual need items, as well as the total number of unmet needs, to similar questions of an established quality-of-life instrument, the World Health Organization Quality of Life-100 (WHOQOL-100)⁴³ (77 participants in Jordan, 79 in Haiti, and 269 in Nepal). For the remaining 11 HESPER items, there was no comparable external criterion available.

Analyses

We performed data analyses with SPSS version 15.0 (SPSS Inc, Chicago, IL). We carried out counts and prevalence rates for categorical demographic variables. We calculated means and standard deviations for continuous demographic variables, time taken to administer the HESPER Scale, time between interviews 1 and 2 (retest), and the number of consistent priority ratings given across raters and time points. We calculated intraclass correlation coefficients (absolute agreement) to assess interrater reliability and test–retest reliability of total number of unmet needs on the HESPER Scale. We calculated percentage agreement and Cohen's κ values to assess interrater and test–retest reliability of

individual HESPER items⁴⁴; we combined “0” (“no serious problem”) and “9” (“not applicable”) ratings into 1 rating for this. Measuring the psychometric properties of individual HESPER items was important, as in humanitarian settings individual item scores are arguably more useful as indicators of perceived needs that can be addressed by aid agencies than the score of the total number of unmet needs.

For criterion (concurrent) validity, we calculated the Pearson’s correlation coefficient to measure the association between total number of unmet needs and total WHOQOL-100 score, and point-biserial correlation coefficients for associations between individual HESPER items and selected questions from the WHOQOL-100. We made predictions for correlation coefficients prior to field-testing, and compared results with these.

RESULTS

We report results separately for the 3 developmental phases of the HESPER Scale.

Phase 1—Development of Draft Scale

The 49 expert survey participants rated all HESPER items as at least moderately important, with means of between 4.88 (SD = 3.27) and 9.39 (SD = 1.15) on a scale of 0 to 10. We therefore took a broad approach in the selection of items into the draft scale for pilot-testing, with the revision of items primarily involving their rephrasing and regrouping. On the basis of participants’ suggestions, we added 1 item (“Health Care”) and also a section to record priority ratings. Overall, we reduced the first draft scale from 38 to 32 items for pilot-testing on the basis of the expert survey.

Furthermore, we conflated the “No Need” and “Met Need” categories of the CANSAS into a single “No Need” (or “No Serious Problem”) category in the draft HESPER Scale. We did this because empirical evidence about moderators^{45,46} and mediators^{47,48} of need indicated that unmet need was most predictive, and also to ease use of the scale in the field.

Phase 2—Pilot-Testing

Cohen’s κ values for interrater reliability of the 32 individual HESPER items included during pilot-testing ranged between 0.62 and

1.0 in Jordan, 0.77 and 1.0 in Gaza, and 0.85 and 1.0 in Sudan. Intraclass correlation coefficients (absolute agreement) for total number of unmet needs were 0.951 in Jordan, 0.998 in Gaza, and 0.998 in Sudan. All items were rated as a serious problem by at least 1 participant in each of the 3 pilot sites. During the participant and interviewer surveys as well as the participant focus group discussions, participants and interviewers indicated that the list of HESPER items was intelligible, comprehensive, culturally acceptable, and useful overall (although suggestions were made for further minor improvements). This established the content validity of the scale.

On the basis of suggested revisions by participants and interviewers during pilot-testing, and on advice from members of the project’s advisory group, we reduced the scale further from 32 to 26 items, primarily by combining closely related items. We reworded parts of the scale to make it more intelligible and restructured it in terms of the order of its items (with basic physical survival needs listed first and items covering community issues last). We also made revisions to training materials.

Phase 3—Field-Testing

Respondents. Participants’ characteristics at field-testing sites are displayed in Table 1. Response rates of people invited to participate were 55.1% in Jordan, 95.0% in Haiti, and 80.0% in Nepal; the response rate across sites was 73.1%. As expected, response rates were relatively low in Jordan, as displaced Iraqi people had previously been exposed to a multitude of surveys and also displayed high levels of fear.

Time to complete. Data collection (330–385 interviews per country) took between 12 and 22 working days (using 12 and 6 interviewers, respectively) in each of the field sites, including time spent on training interviewers.

On average, the HESPER Scale took 14.8 (SD = 4.1) minutes to complete in Jordan, 21.3 (SD = 11.5) minutes in Haiti, and 22.0 (SD = 6.0) minutes in Nepal; across sites, the mean was 19.5 minutes (SD = 8.7).

Interrater reliability. Intraclass correlation coefficients (absolute agreement) for interrater reliability of total number of unmet needs were 0.998 in Jordan, 0.986 in Haiti, and 0.995

in Nepal; across sites it was 0.998. Percentage agreements for interrater reliability of need ratings of individual HESPER items ranged between 95.3% and 100%, and Cohen’s κ ranged between 0.66 and 1.0 across the 3 field-testing sites (Table 2).

The mean number of priority ratings that raters agreed on was 3.0 (SD = 0) in Jordan, 3.0 (SD = 0) in Haiti, and 2.95 (SD = 0.22) in Nepal; across sites it was 2.98 (SD = 0.12) (out of 3.0).

Test–retest reliability. Retest interviews were conducted between 6 and 8 days following the first interview in Jordan, and between 5 and 8 days later in Nepal; the means were 6.9 days (SD = 0.3) and 6.5 days (SD = 0.8), respectively.

Intraclass correlation coefficients (absolute agreement) for test–retest reliability of total number of unmet needs were 0.961 in Jordan and 0.773 in Nepal; across the 2 sites it was 0.907. Percentage agreements for test–retest reliability of need ratings of individual HESPER items ranged between 66.7% and 100%, and Cohen’s κ ranged between 0.07 and 1.0 across the 2 sites (Table 3).

The mean number of priority ratings that were consistently given at the 2 time points were 2.4 (SD = 0.71) in Jordan and 1.33 (SD = 0.79) in Nepal; across the 2 sites, the mean was 1.86 (SD = 0.92; out of 3.0).

As test-retest reliability results in Nepal were lower overall than all other reliability results across the 3 field sites, brief interviews were conducted with 12 participants following retest interviews in Nepal, where they were asked for reasons why they may have responded differently at interviews 1 and 2. Reasons given included the following:

- They believed the collaborating agency would be more likely to offer them support if they mentioned a wide range of different problems during the 2 interviews (n = 7).
- They had been experiencing some tensions in one of the interviews, for instance because family members had been resettled (n = 5).
- They were old or had low levels of understanding or listening skills (n = 3).
- Discussions with family members following the first interview led them to respond differently during the second interview (n = 3).

TABLE 1—Demographic Characteristics of Study Participants at the 3 Field-Testing Sites for the Humanitarian Emergency Settings Perceived Needs (HESPER) Scale: 2010

Characteristic	Total (n = 817), No. (%) or Mean (SD)	Jordan (n = 269), No. (%) or Mean (SD)	Haiti (n = 279), No. (%) or Mean (SD)	Nepal (n = 269), No. (%) or Mean (SD)
Gender				
Male	305 (37.3)	116 (43.1)	50 (17.9)	139 (51.7)
Female	512 (62.7)	153 (56.9)	229 (82.1)	130 (48.3)
Age, y	37.09 ±13.5	40.24 ±13.36	34.22 ±12.31	36.92 ±14.15
Marital status				
Married	441 (54.0)	191 (71.0)	33 (12.0)	217 (80.7)
Unmarried	335 (41.0)	56 (20.8)	229 (82.1)	50 (18.6)
Widowed	18 (2.2)	16 (5.9)	2 (0.7)	0
Divorced or separated	8 (1.0)	6 (2.2)	0	2 (0.7)
Cohabiting	11 (1.3)	0	11 (4.0)	0
No. of children	2.37 ±2.17	2.11 ±1.95	2.59 ±2.14	2.39 ±2.37
Level of education				
Illiterate or no formal education	164 (20.1)	7 (2.6)	49 (17.7)	108 (40.1)
Primary school (grades 1–6)	190 (23.3)	29 (10.8)	98 (35.4)	63 (23.4)
Secondary school (grades 7–12)	315 (38.6)	104 (38.7)	122 (44.0)	89 (33.1)
University	146 (17.9)	129 (48.0)	8 (2.9)	9 (3.3)
Religion				
Christian	329 (40.3)	45 (16.7)	268 (96.1)	16 (5.9)
Muslim	221 (27.1)	221 (82.2)	0	0
Hindu	178 (21.8)	0	0	178 (66.2)
Buddhist	52 (6.4)	0	0	52 (19.3)
Other religion ^a	27 (3.3)	3 (1.1)	1 (0.4)	23 (8.6)
No religion	5 (0.6)	0	5 (1.8)	0
Time displaced, y	7.77 ±8.09	3.84 ±2.18	0.67 ±0.06	18.95 ±0.93

Note. Numbers do not always add up to total score because of missing data.

^aOther religions include Kirat, Sanatan, Biswasi, Manab, Nastak (Nepal), Haba'i, Sa'aebiya (Jordan), and Voodoo (Haiti).

Criterion (concurrent) validity. Total number of unmet needs on the HESPER Scale correlated with the total WHOQOL-100 score as was predicted before data collection (i.e., Pearson's correlation was within 1 order-of-magnitude step of the predicted value, where 0.1–0.3 represented a low correlation, 0.3–0.5 represented a medium correlation, and 0.5–1.0 represented a high correlation) in all 3 settings ($r = -0.629$ in Jordan, -0.417 in Haiti, and -0.469 in Nepal), as well as with the WHOQOL-100 question "How would you rate your quality of life?" ($r = -0.501$ in Jordan, -0.302 in Haiti, and -0.286 in Nepal).

Point-biserial correlations between 15 of the 26 individual HESPER items and 25 related WHOQOL-100 questions were also mostly as was predicted before data collection in all 3

field sites, apart from the item "Income or Livelihood" in Haiti ($r = 0.033$ and 0.242 for 2 related WHOQOL-100 questions, where negative low to medium and negative low correlations had been predicted, respectively), the item "Distress" in Haiti ($r = 0.06$ and 0.078 , where negative low and positive medium correlations had been predicted, respectively), the item "The Way Aid Is Provided" in Nepal ($r = 0.015$, where a negative low correlation had been predicted), and the item "Safety or Protection From Violence for Women in Your Community" in Nepal ($r = 0.045$, where a negative low correlation had been predicted). In Haiti, however, validation for the 2 items was compromised, as the items were rated as serious problems by over 90% of participants (i.e., limited variability and power).

Finalization of HESPER Scale. We made minor changes in the wording of 8 items to finalize the HESPER Scale following field-testing; for example, the item heading "Aid" was rephrased as "The Way Aid Is Provided," and for the item "Clothing, Shoes, Bedding or Blankets" the word "Clothing" was replaced with "Clothes."

DISCUSSION

The HESPER Scale proved to be a valuable and comprehensive tool, with adequate psychometric properties across different population groups in a variety of humanitarian settings. Interrater and test–retest reliability results were good to very good overall. International experts, as well as interviewers and participants in several pilot sites, found

TABLE 2—Percentage Agreement and Cohen's κ for Interrater Reliability of Need Ratings of Individual Humanitarian Emergency Settings Perceived Needs (HESPER) Scale Items

HESPER Items	Total (n = 132), Cohen's κ (% Agreement)	Jordan ^a (n = 46), Cohen's κ (% Agreement)	Haiti ^b (n = 44), Cohen's κ (% Agreement)	Nepal (n = 42), Cohen's κ (% Agreement)
Drinking water	0.98 (99.2)	1.0 (100)	0.94 (97.7)	1.0 (100)
Food	0.97 (98.5)	0.94 (97.8)	0.79 (97.7)	1.0 (100)
Place to live in	0.98 (99.2)	0.96 (97.8)	(100) ^c	1.0 (100)
Toilets	0.95 (97.7)	0.94 (97.8)	0.89 (95.3)	1.0 (100)
Keeping clean	0.99 (99.2)	1.0 (100)	1.0 (100)	0.95 (97.6)
Clothing, shoes, bedding, or blankets	0.98 (99.2)	1.0 (100)	1.0 (100)	0.95 (97.6)
Income or livelihood	1.0 (100)	1.0 (100)	1.0 (100)	1.0 (100)
Physical health	0.97 (98.5)	1.0 (100)	0.95 (97.7)	0.95 (97.6)
Health care	0.95 (97.7)	0.96 (97.8)	0.88 (95.5)	1.0 (100)
Distress	1.0 (100)	1.0 (100)	1.0 (100)	1.0 (100)
Safety	1.0 (100)	1.0 (100)	1.0 (100)	1.0 (100)
Education for your children	0.97 (98.5)	0.91 (97.8)	0.94 (97.7)	1.0 (100)
Care for family members	0.94 (97.0)	0.9 (95.7)	(95.5) ^c	1.0 (100)
Support from others	1.0 (100)	1.0 (100)	1.0 (100)	1.0 (100)
Separation from family members	1.0 (100)	1.0 (100)	1.0 (100)	1.0 (100)
Being displaced from home	1.0 (100)	1.0 (100)	1.0 (100)	1.0 (100)
Information	0.97 (98.5)	1.0 (100)	0.66 (97.7)	0.93 (97.6)
Aid	0.98 (99.2)	0.95 (97.8)	1.0 (100)	1.0 (100)
Respect	0.98 (99.2)	0.9 (97.8)	1.0 (100)	1.0 (100)
Moving between places	0.95 (97.7)	0.95 (97.8)	0.89 (95.5)	1.0 (100)
Too much free time	0.98 (99.2)	1.0 (100)	1.0 (100)	0.94 (97.6)
Law and justice in your community	0.99 (99.2)	1.0 (100)	0.92 (97.7)	1.0 (100)
Safety or protection from violence for women in your community	0.95 (97.7)	0.9 (97.8)	0.89 (95.5)	1.0 (100)
Alcohol or drug use in your community	0.98 (99.2)	1.0 (100)	0.94 (97.7)	1.0 (100)
Mental illness in your community	0.97 (98.5)	1.0 (100)	0.91 (95.5)	1.0 (100)
Care for people in your community who are on their own	1.0 (100)	1.0 (100)	1.0 (100)	1.0 (100)

Note. "0" ("No Serious Problem") and "9" ("Not Applicable") ratings have been combined.

^aIn Jordan, an additional item, "Residency or Resettlement," was added on the basis of findings made during pilot-testing (percentage agreement = 100, Cohen's κ = 1.0).

^bIn Haiti, an additional item, "Burying and mourning the dead in your community," was added on the basis of field observations (percentage agreement = 97.7, Cohen's κ = 0.94).

^cNot possible to compute Cohen's κ , as ratings for at least 1 of the variables was a constant.

the list of HESPER items to be comprehensive and relevant, providing evidence for content validity of the scale. Furthermore, most HESPER items correlated with related questions of the WHOQOL-100 as was predicted before data collection, suggesting criterion validity.

Limitations

Because of issues of feasibility, there were some limitations in the way the HESPER Scale's psychometric properties were measured. In particular, the method of having a second interviewer silently rate the HESPER Scale to assess interrater reliability may overestimate interrater reliability, as responses may

be affected by the personal characteristics and manner of interviewers. Moreover, although the WHOQOL has been widely used and validated worldwide,⁴⁹ it had not been validated in the populations in which the HESPER Scale was field-tested, thereby reducing the strength of the assessed validity.

Whereas interrater reliability across the 3 field-testing sites and test–retest reliability in Jordan was excellent, test–retest reliability in Nepal was substantially lower. Ten of 12 participants in Nepal who were asked to provide an explanation for this indicated that they made some deliberate effort to respond differently during the 2 interviews. This suggests reduced validity of the retest results in Nepal,

as it may be a reflection of affected populations' conscious attempts to influence humanitarian response (for instance, by overestimating the seriousness of their needs).⁵⁰ Although the psychometric results so far are very promising, these issues highlight the need for more work to be conducted across different settings, to provide further evidence for reliability and validity of the HESPER Scale. It may be useful for this to include an assessment of construct validity and internal consistency of the scale, in particular when working with total scores of unmet need. Furthermore, factor or principal component analyses may be valuable in identifying underlying structures of associated HESPER items.

TABLE 3—Percentage Agreement and Cohen's κ for Test–Retest Reliability of Need Ratings of Individual Humanitarian Emergency Settings Perceived Needs (HESPER) Scale Items

HESPER Items	Total (n = 122), Cohen's κ (% Agreement)	Jordan ^a (n = 59), Cohen's κ (% Agreement)	Nepal (n = 63), Cohen's κ (% Agreement)
Drinking water	0.82 (91.7)	0.89 (94.9)	0.17 (88.7)
Food	0.66 (82.8)	0.9 (94.9)	0.43 (71.4)
Place to live in	0.66 (82.8)	0.86 (93.2)	0.43 (73.0)
Toilets	0.63 (85.2)	0.88 (94.9)	0.39 (76.2)
Keeping clean	0.64 (84.4)	0.73 (88.1)	0.55 (81.0)
Clothing, shoes, bedding or blankets	0.67 (83.6)	0.93 (96.6)	0.43 (71.4)
Income or livelihood	0.73 (91.8)	1.0 (100)	0.6 (84.1)
Physical health	0.6 (80.2)	0.77 (89.8)	0.38 (71.0)
Health care	0.75 (87.7)	0.8 (91.5)	0.49 (84.1)
Distress	0.7 (85.2)	0.81 (94.9)	0.39 (76.2)
Safety	0.56 (85.2)	0.71 (89.8)	0.42 (81.0)
Education for your children	0.71 (93.4)	0.88 (96.6)	0.46 (90.5)
Care for family members	0.69 (86.0)	0.89 (94.9)	0.45 (77.4)
Support from others	0.85 (93.4)	0.86 (93.2)	0.47 (93.7)
Separation from family members	0.68 (85.2)	0.86 (96.6)	0.49 (74.6)
Being displaced from home	0.65 (86.8)	1.0 (100)	0.48 (74.2)
Information	0.52 (79.5)	0.69 (84.7)	0.07 (74.6)
Aid	0.75 (87.7)	0.84 (94.9)	0.38 (81.0)
Respect	0.76 (91.8)	0.84 (93.2)	0.61 (90.5)
Moving between places	0.64 (85.2)	0.85 (93.2)	0.39 (77.8)
Too much free time	0.59 (79.5)	0.86 (93.2)	0.26 (66.7)
Law and justice in your community	0.55 (82.0)	0.66 (86.4)	0.46 (77.8)
Safety or protection from violence for women in your community	0.62 (87.7)	0.77 (94.9)	0.52 (81.0)
Alcohol or drug use in your community	0.67 (88.5)	0.79 (98.3)	0.57 (79.4)
Mental illness in your community	0.79 (90.2)	0.83 (91.5)	0.65 (88.9)
Care for people in your community who are on their own	0.64 (82.8)	0.76 (88.1)	0.51 (77.8)

Note. Participants with a change in their condition were excluded from the analyses. "0" ("No Serious Problem") and "9" ("Not Applicable") ratings have been combined. Test–retest reliability was not measured in Haiti, as it was not considered appropriate in this setting.

^aIn Jordan, an additional item, "Residency or Resettlement," was added on the basis of findings made during pilot-testing (percentage agreement = 96.6, Cohen's κ = 0.92).

Sampling methods were often challenging. In particular, as there was no complete list of households or individuals available in Jordan and Haiti during field-testing, random-walk methods had to be employed. Furthermore, the response rate in Jordan was relatively low. The findings may therefore not be representative of the affected populations at large in the 3 settings. However, the effect of such biases on psychometric estimates is likely to be minimal, as the focus is more on substantive responses than on the representativeness of participants.

Implications

The HESPER Scale enables the perceived problems of people living in humanitarian

situations to be assessed quickly and reliably, directly on the basis of their own views. The scale has been found to be applicable and useful in several diverse humanitarian settings, and is available in English, French, Spanish, Arabic, Nepali, and Haitian Creole. So far, the HESPER Scale has been tested only in adult populations.

However, use of the HESPER Scale at one time point is not sufficient to understand the complexities of population needs. Needs assessments should be viewed and contextualized within the specific timeframe within which they are conducted; for this purpose, it may be that the HESPER Scale can be used repeatedly over time to identify shifts and trends

in perceived needs and to assess whether needs are being addressed adequately over time. To assess this possibility, it would be useful for future research to measure the scale's sensitivity to change, something that was beyond the scope of the current study.

Moreover, the HESPER Scale on its own may not be sufficient to fully understand people's perceived needs, nor will it directly indicate what is required to respond to these needs. HESPER surveys can be followed up with in-depth key informant interviews to better understand the specifics of why—from the participants' perspectives—needs are rated as they are. There is a continued need for traditional surveillance and early warning

systems to identify needs. The HESPER Scale is not able to function as an operational tool to give detailed feedback on the quality of interventions within sectors. However, in situations where interventions have started to respond to needs and affected populations still indicate that a particular issue ranks high as need, the HESPER Scale may give a strong indication that the response does not yet meet these needs.

Conclusions

The development of the HESPER Scale opens up new avenues in the science of humanitarian needs assessment by (1) enabling rapid representative mapping and ranking of perceived needs as expressed by affected populations (allowing for differentiation of perceived needs between different population subgroups) and (2) showing that not just psychopathology¹⁵ but also the broad spectrum of humanitarian needs can be assessed with documented reliability and validity. It offers a method to produce information that can be directly used to prioritize and guide specific forms of emergency relief and to assess the impact of their implementation. This type of assessment allows affected populations to express what they consider to be their needs. The HESPER Scale thereby fills a gap within the multisectoral needs assessment field, allowing comparisons to be made between the views of international aid agencies and affected populations of what is needed, and therefore facilitating priorities for the most appropriate humanitarian response to be set. ■

About the Authors

Maya Semrau, Louise M. Howard, Mike Slade, and Graham Thornicroft are with the Health Service and Population Research Department, Institute of Psychiatry, King's College London, London, UK. Mark van Ommeren is with the Department of Mental Health and Substance Abuse, World Health Organization, Geneva, Switzerland. Monica Blagescu is with Humanitarian Accountability Partnership (HAP) International, Abuja, Nigeria. Andre Griekspoor is with Health Action in Crises Cluster, World Health Organization, Geneva, Switzerland. Mark Jordans is with HealthNet TPO, Amsterdam, The Netherlands. Heidi Lempp is with the School of Medicine, Academic Rheumatology, King's College London. At the time of the study, Anita Marini was with the World Health Organization, Jordan Country Office, Amman, Jordan. Jon Pedersen is with Fafo Institute for Applied International Studies, Oslo, Norway. At the time of the study, Isabelle Pilotte was with International Medical Corps, Port-au-Prince, Haiti.

Correspondence should be sent to Graham Thornicroft, Institute of Psychiatry, Box P029, De Crespigny Park,

London SE5 8AF, UK (e-mail: graham.thornicroft@kcl.ac.uk). Reprints can be ordered at <http://www.ajph.org> by clicking the "Reprints" link.

This article was accepted January 30, 2012.

Note. The views expressed in this report are those of the authors only and do not necessarily represent the decisions, policies, or views of their employers.

Contributors

M. van Ommeren and G. Thornicroft jointly conceptualized development of the HESPER Scale. M. Semrau and M. van Ommeren led the HESPER project. M. Semrau conducted each of the 7 pilot and field studies in close consultation with M. van Ommeren, performed the data analyses, and led the writing of the article; M. van Ommeren provided technical advice and detailed input throughout. G. Thornicroft and L. M. Howard supervised M. Semrau and gave technical advice. H. Lempp and A. Griekspoor are members of the HESPER steering committee. M. Slade informed the design of the HESPER Scale. J. Pedersen and M. Jordans facilitated data collection in Gaza and Nepal, respectively. I. Pilotte, A. Marini, and M. Blagescu coordinated data collection in Haiti, Jordan, and Sudan, respectively. All authors contributed to the writing of the article, and have seen and approved the final version.

Acknowledgments

Maya Semrau is funded by a PhD studentship grant of the Medical Research Council (UK). Graham Thornicroft is funded by a NIHR Applied Programme grant awarded to the South London and Maudsley NHS Foundation Trust, and is affiliated with the NIHR Specialist Mental Health Biomedical Research Centre at the Institute of Psychiatry, King's College London and the South London and Maudsley NHS Foundation Trust. Louise M. Howard is supported by an NIHR Programme Grant for Applied Research (RP-PG-0108-10084) and by the NIHR Specialist Mental Health Biomedical Research Centre at the Institute of Psychiatry, King's College London, and the South London and Maudsley NHS Foundation Trust. Heidi Lempp receives half of her salary from Guy's and St. Thomas' Charity, London.

The development of the HESPER Scale was a collaborative project between the Department of Mental Health and Substance Abuse at WHO and the Institute of Psychiatry at King's College London. The steering group of the HESPER project included Mark van Ommeren and Andre Griekspoor at WHO Geneva and Graham Thornicroft, Louise M. Howard, Heidi Lempp, Morven Leese, and Maya Semrau at King's College London.

The HESPER international advisory group consisted of Paul Bolton (John Hopkins Bloomberg School of Public Health), Kaz de Jong (Medicins Sans Frontieres Amsterdam), Nadine Ezard (Monash University), Richard Garfield (Columbia University), Johan Heffinck (DG ECHO, European Commission), Lynne Jones (International Medical Corps), Helen McColl (International Rehabilitation Council for Torture Victims), Pau Pérez-Sales (Medicos del Mundo, MdM-E), Shekhar Saxena (WHO), Mike Slade (King's College London), Egbert Sondorp (London School of Hygiene and Tropical Medicine), Zachary Steel (University of New South Wales), Wietse Tol (HealthNet TPO, Yale University), and Mike Wessells (Columbia University, New York).

Data collection in Gaza was organized by Fafo Institute for Applied International Studies, with the WHO Office in Gaza providing advice and funding provided by WHO Geneva and Fafo. Data collection in Haiti was

organized by International Medical Corps Haiti and was funded through a PhD studentship grant by the Medical Research Council (UK) to Maya Semrau. Data collection in Jordan was organized by WHO Jordan and was implemented by Accurate Opinion (field-testing) and the Market Research Organisation (pilot-testing); UNHCR provided advice on sampling. Data collection in Jordan was funded by the Jordanian Nursing Council, WHO Jordan, and the University of London Central Research Fund. Data collection in Nepal was organized by HealthNet TPO/TPO Nepal, with funding provided by WHO Geneva, and UNHCR Nepal and WHO Nepal providing further support. Data collection in Sudan was organized by Humanitarian Accountability Partnership (HAP) International. Data collection in the United Kingdom was facilitated by the British Refugee Council.

Note. The funding sources had no involvement in the design or execution of the study, nor in the data analyses or the decision to submit the article for publication.

Human Participant Protection

Ethics approval for the study was obtained through the King's College London Psychiatry, Nursing and Midwifery Research Ethics Committee. In Nepal, further ethical approval was obtained from the Nepal Health Research Council and in Jordan further permission for the study was obtained from the Ministry of Interior, Ministry of Planning, and Ministry of Health. Participants in all 3 phases gave their free written or verbal consent to take part.

References

1. IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva, Switzerland: Inter-Agency Standing Committee; 2007.
2. Oxfam GB for the Emergency Capacity Building Project. *Impact Measurement and Accountability in Emergencies—The Good Enough Guide*. Oxford, UK: Oxfam GB; 2007.
3. *Participation of Crisis-Affected Populations in Humanitarian Action: A Handbook for Practitioners*. London, UK: Active Learning Network for Accountability and Performance in Humanitarian Action; 2003.
4. *HAP 2007 Standard in Humanitarian Accountability and Quality Management*. Geneva, Switzerland: Humanitarian Accountability Partnership International (HAP); 2007.
5. *Humanitarian Charter and Minimum Standards in Disaster Response*. Geneva, Switzerland: Sphere Project; 2004.
6. *Humanitarian Charter and Minimum Standards in Disaster Response*. Geneva, Switzerland: Sphere Project; 2011.
7. Tol WA, Patel V, Tomlinson M, et al. Research priorities for mental health and psychosocial support in humanitarian settings. *PLoS Med*. 2011;8(9):e1001096.
8. Pérez-Sales P, Cervellón P, Vázquez C, Vidales D, Gaborit M. Post-traumatic factors and resilience: the role of shelter management and survivors' attitudes after the earthquakes in El Salvador (2001). *J Community Appl Soc Psychol*. 2005;15(5):368–382.
9. Weiss WM, Bolton P, Shankar AV. Rapid Assessment Procedures (RAP): addressing the perceived needs of refugees and internally displaced persons through participatory learning and action. 2nd ed. Center for Refugee and Disaster Studies. 2000. Available at:

- http://www.jhsph.edu/refugee/publications_tools/publications/rap.html. Accessed June 20, 2012.
10. Miller KE, Rasmussen A. War exposure, daily stressors, and mental health in conflict and post-conflict settings: bridging the divide between trauma-focused and psychosocial frameworks. *Soc Sci Med*. 2010;70(1):7–16.
 11. Bolton P, Tang AM. Using ethnographic methods in the selection of post-disaster, mental-health interventions. *Prehosp Disaster Med*. 2004;19(1):97–101.
 12. HNTS-SISN. Health and Nutrition Tracking Service. 2011. Available at: <http://www.thehnts.org>. Accessed June 20, 2012.
 13. Standardized Monitoring and Assessment of Relief and Transitions. 2011. Available at: <http://www.smartindicators.org/index.html>. Accessed June 20, 2012.
 14. *Emergency Food Security Assessment Handbook*. 2nd ed. Rome, Italy: World Food Programme; 2009.
 15. Hollifield M, Warner TD, Lian N, et al. Measuring trauma and health status in refugees: a critical review. *JAMA*. 2002;288(5):611–621.
 16. Jordans MJD, Komproe IH, Tol WA, de Jong JTVM. Screening for psychosocial distress amongst war-affected children: cross-cultural construct validity of the CPDS. *J Child Psychol Psychiatry*. 2009;50(4):514–523.
 17. Bolton P, Tang AM. An alternative approach to cross-cultural function assessment. *Soc Psychiatry Psychiatr Epidemiol*. 2002;37(11):537–543.
 18. World Health Organization (WHO) and King's College London. *The Humanitarian Emergency Settings Perceived Needs Scale (HESPER): Manual With Scale*. Geneva, Switzerland: WHO; 2011. Available at: http://whqlibdoc.who.int/publications/2011/9789241548236_eng.pdf. Accessed October 24, 2011.
 19. Slade M, Thornicroft G, Loftus L, Phelan M, Wykes T. *CAN: Camberwell Assessment of Need—A Comprehensive Needs Assessment Tool for People With Severe Mental Illness*. London, UK: Gaskell; 1999.
 20. Phelan M, Slade M, Thornicroft G, et al. The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. *Br J Psychiatry*. 1995;167(5):589–595.
 21. Andresen R, Caputi P, Oades LG. Interrater reliability of the Camberwell Assessment of Need Short Appraisal Schedule. *Aust N Z J Psychiatry*. 2000;34(5):856–861.
 22. Xenitidis K, Thornicroft G, Leese M, et al. Reliability and validity of the CANDID—a needs assessment instrument for adults with learning disabilities and mental health problems. *Br J Psychiatry*. 2000;176:473–478.
 23. Reynolds T, Thornicroft G, Abas M, et al. Camberwell Assessment of Need for the Elderly (CANE)—development, validity and reliability. *Br J Psychiatry*. 2000;176:444–452.
 24. Howard L, Hunt K, Slade M, et al. Assessing the needs of pregnant women and mothers with severe mental illness: the psychometric properties of the Camberwell Assessment of Need—Mothers (CAN-M). *Int J Methods Psychiatr Res*. 2007;16(4):177–185.
 25. Thomas SDM, Slade M, McCrone P, et al. The reliability and validity of the Forensic Camberwell Assessment of Need (CANFOR): a needs assessment for forensic mental health service users. *Int J Methods Psychiatr Res*. 2008;17(2):111–120.
 26. McCrone P, Leese M, Thornicroft G, et al. Reliability of the Camberwell Assessment of Need—European Version. EPSILON Study 6. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. *Br J Psychiatry Suppl*. 2000;(39):s34–s40.
 27. McColl H, Johnson S. Characteristics and needs of asylum seekers and refugees in contact with London community mental health teams: a descriptive investigation. *Soc Psychiatry Psychiatr Epidemiol*. 2006;41(10):789–795.
 28. McCrone P, Bhui K, Craig T, et al. Mental health needs, service use and costs among Somali refugees in the UK. *Acta Psychiatr Scand*. 2005;111(5):351–357.
 29. Semrau M. *The HESPER (Humanitarian Emergency Settings Perceived Needs) Scale—Development of a Draft Instrument* [MSc Dissertation]. London, UK: Institute of Psychiatry, King's College London.
 30. Lee K, Bolton P. Impact of FilmAid programs in Kakuma, Kenya—final report. 2007. Available at: http://libcloud.s3.amazonaws.com/129/43/1/437/BU_REPORT_FINAL_EVALUATION.pdf. Accessed June 20, 2012.
 31. Bolton P, Ndogoni L. Cross-cultural assessment of trauma-related mental illness: CERTI Crisis and Transition Toolkit. 2000. Available at: <http://www.certigroup.com/publications/Manuals/cross-cultural-10.PDF>. Accessed June 20, 2012.
 32. Fritz Institute. The immediate response to the Java Tsunami: perceptions of the affected. 2007. Available at: www.fritzinstitute.org. Accessed June 20, 2012.
 33. Barton T, Mutiti A, The Assessment Team for Psycho-Social Programmes in Northern Uganda. *NUPSNA—Northern Uganda Psycho-Social Needs Assessment*. Kisubi, Uganda: UNICEF/Uganda Government; 1998.
 34. Betancourt TS, Speelman L, Onyango G, Bolton P. A qualitative study of mental health problems among children displaced by war in northern Uganda. *Transcult Psychiatry*. 2009;46(2):238–256.
 35. Fritz Institute. Recipient perceptions of aid effectiveness: rescue, relief and rehabilitation in tsunami affected Indonesia, India and Sri Lanka. 2005. Available at: www.fritzinstitute.org. Accessed June 20, 2012.
 36. Fritz Institute. Recovering from the Java Earthquake: perceptions of the affected. 2007. Available at: www.fritzinstitute.org. Accessed June 20, 2012.
 37. Fritz Institute. Surviving the Pakistan Earthquake: perceptions of the affected one year later. 2006. Available at: www.fritzinstitute.org. Accessed June 20, 2012.
 38. Murray L, Bass J, Bolton P. Qualitative study to identify indicators of psychosocial problems and functional impairment among residents of Sange District, South Kivu, Eastern DRC. 2006. Available at: http://pdf.usaid.gov/pdf_docs/PNADI610.pdf. Accessed June 20, 2012.
 39. van Ommeren M, Sharma B, Thapa S, et al. Preparing instruments for transcultural research: use of the translation monitoring form with Nepali-speaking Bhutanese refugees. *Transcult Psychiatry*. 1999;36(3):285–301.
 40. World Health Organization. Process of translation and adaptation of instruments. Available at: http://www.who.int/substance_abuse/research_tools/translation/en. Accessed October 19, 2011.
 41. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;17:101–110.
 42. Kish L. A procedure for objective respondent selection within the household. *J Am Stat Assoc*. 1949;44(247):380–387.
 43. World Health Organization (WHO), WHOQOL-Group, Division of Mental Health. *WHOQOL-100*. Geneva, Switzerland: WHO; 1995.
 44. Streiner DL, Norman GR. *Health Measurement Scales—A Practical Guide to Their Development and Use*. 4th ed. Oxford, UK: Oxford University Press; 2008.
 45. Ruggeri M, Leese M, Slade M, Bonizzato P, Fontecedro M, Tansella M. Demographic, clinical, social and service variables associated with higher needs for care in community psychiatric service patients. The South Verona Outcome Project 8. *Soc Psychiatry Psychiatr Epidemiol*. 2004;39(1):60–68.
 46. Junghan U, Leese M, Priebe S, Slade M. Staff and patient perspectives on unmet need and therapeutic alliance in community services. *Br J Psychiatry*. 2007;191:543–547.
 47. Slade M, Leese M, Ruggeri M, Kuipers E, Tansella M, Thornicroft G. Does meeting needs improve quality of life? *Psychother Psychosom*. 2004;73(3):183–189.
 48. Slade M, Leese M, Cahill S, Thornicroft G, Kuipers E. Patient-rated mental health needs and quality of life improvement. *Br J Psychiatry*. 2005;187:256–261.
 49. Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF Quality of Life Assessment: psychometric properties and results of the international field trial. A Report from the WHOQOL Group. *Qual Life Res*. 2004;13(2):299–310.
 50. Bailey S. *Need and Greed: Corruption, Risks, Perceptions and Prevention in Humanitarian Assistance—HPG Policy Brief 32*. London, UK: London Overseas Development Institute; 2008.