DEPARTMENT OF HEALTH

Minnesota Refugee Mental Health Screening Tool Final Report

PILOT PROJECT METHODS, RESULTS, AND RECOMMENDATIONS

12/30/2021

Minnesota Department of Health
Refugee Health Program
PO Box 64975
St. Paul, MN 55164-0975
refugeehealth@state.mn.us
www.health.state.mn.us

To obtain this information in a different format, call: 651-201-5827.

Contents

Acknowledgements

The Minnesota Department of Health Refugee Health Program would like to thank the many individuals and organizations who helped us develop the Minnesota Mental Health Screening Tool and shaped the implementation of the pilot. We would particularly like to acknowledge the pilot sites: HealthPartners Center for International Health, the Hennepin County Public Health Clinic, the Mayo-Rochester Health Care System, M Health Fairview Roselawn, and Olmsted County Public Health. We also thank the Center for Victims of Torture and Dr. Patricia Shannon of the University of Minnesota-Twin Cities School of Social Work. Finally, we acknowledge the work of Dr. Maria Vukovich who conducted the Cross-Cultural Measurement Analysis.

We would also like to acknowledge the Centers for Disease Control and Prevention for their support of the Minnesota Center of Excellence in Refugee Health (5 NU50CK000459); their support helped us complete the pilot.

Pilot partners

Jaya Durvasula, MD
HealthPartners Center for International Health

Leah Espinda-Brandt, RN
Olmsted County Public Health

Margaret Fitzthum, CNP
Hennepin County Public Health Clinic

Ellen Frerich, MN, MSW, MPP
Minnesota Department of Health
Refugee Health Program

Mateo Frumholtz
Minnesota Department of Health
Refugee Health Program

Ahmed Hassan, MA, LPCC
Summit Guidance

Laura Heeter, MD, MPH
HealthPartners Center for International Health

Marge Higgins, LSW
Minnesota Department of Health
Refugee Health Program

Marisa Hynes, RN
Olmsted County Public Health

Kelley Jewett, MD
M Health Fairview Roselawn Clinic

Georgi Kroupin, MA, LP
HealthPartners Center for International Health

Jim Letts, MD
M Health Fairview Roselawn Clinic

Blain Mamo, MPH
Minnesota Department of Health
Refugee Health Program

Christine May, CNP
Hennepin County Public Health Clinic

Mary Jo Meuleners
Hennepin County Public Health Clinic

Andrea Northwood, PhD, LP
The Center for Victims of Torture

Eliza Priest, CNP
Hennepin County Public Health Clinic
Refugee Health Program

Nasreen S. Quadri, MD
HealthPartners Center for International Health

Ann Settgast, MD, DTM&H
HealthPartners Center for International Health

Patricia Shannon, PhD, LP
University of Minnesota, School of Social Work

Conard Simba, RN
Hennepin County Public Health Clinic

Shana Sniffen, MD
M Health Fairview Roselawn Clinic

Eh Taw Dwe
The Center for Victims of Torture

Larissa Turin, RN
HealthPartners Center for International Health

Kailey Urban, MPH
Minnesota Department of Health
Refugee Health Program

Lori Wald, RN
St. Paul-Ramsey County Department of Public Health, Refugee Health Program

Patricia Walker, MD, DTM&H
University of Minnesota / Division of Infectious Disease & International Medicine, Department of Medicine, HealthPartners Center for International Health

Maria Vukovich, PhD
The Center for Victims of Torture

Executive summary

This document summarizes the pilot of the Minnesota Refugee Mental Health Screening Tool, the findings of that pilot, and the finalization of the Minnesota Refugee Mental Health Screening Tool. The pilot for this tool consisted of two versions implemented at four pilot clinics within the state of Minnesota, from January 1, 2016 – January 31, 2020. Overall, approximately 13% of 1,672 adult refugee arrivals screened positive for mental health distress using the Minnesota Refugee Mental Health Screening Tool (either version), with differing prevalence by country of origin and age that is roughly equivalent to that observed in other studies. Cross-cultural measurement analyses were conducted to explore appropriateness of the items in the two versions in assessing mental health distress in newcomers from the five countries most represented in resettlement in Minnesota. Based on these analyses, the final adapted screening instrument incorporates items from each version.

Psychometrically sound and cross-culturally equivalent screening instruments are essential to accurately identify the extent to which traumatic experiences and resulting mental health distress affect the well-being of refugees and to what extent endorsement of mental health distress may differ across specific cultural groups. A measurement analyses of two brief measures developed to screen for mental health distress in recently resettled adult refugees was performed to evaluate the cross-cultural utility and equivalence across six distinct countries of origin: Burma, Bhutan, Ethiopia, Iraq, Somalia, and the Democratic Republic of

Congo (DR Congo). The aim was to identify a final version of a brief mental health screening instrument that performed well among refugees from all countries of origin.

Results of the measurement analyses indicated the mean level of reported mental health distress was greater for refugees from Iraq, the DR Congo, and Bhutan than mean level of mental health distress reported by refugees from Burma, Ethiopia, and Somalia on both measures. These findings are consistent with previous literature that has found elevated rates of depression, anxiety and PTSD reported among those from Iraq, Bhutan, and the DR Congo as compared to other refugee populations in the project. Although some mean differences in reported mental health distress were found across refugees from six countries of origin, the average fit of persons was found to be appropriate for refugees overall. Fit of persons was poorer for refugees from Bhutan and Iraq on both measures, likely a function of smaller sample sizes (<100) for these groups.

Rates of endorsement on items differed by country of origin. Items are listed in Table 2 below. Overall, **item equivalence across refugee countries of origin differed on three key items: avoidance (AVD-4) on the Version A tool, and dreams (DRE-9) and memory (MEM-10) on the Version B tool**. After the AVD-4 item was removed, the fit statistics and reliability of the Version A tool were reduced. Removal of the MEM-10 item on the Version B also resulted in poorer fit and reliability of the measure. The findings confirm the avoidance and memory items should be retained as they contributed to accurately identifying mental health distress in refugees across countries of origin. Version B did not worsen in fit statistics or reliability when DRE-9 was removed. Thus, it is recommended this item be removed from the scale or merged with the sleep (SLE-8) item.

Items included in the final mental health screening tool are detailed below in Table 1. Expanded data collection through state-wide implementation of the final mental health screening instrument in parallel with validated diagnostic instruments is a suggested next step. This effort would allow for confirmation of whether current screening items are valid and predictive of mental health distress in resettled refugee populations.

Table 1. Final Version of Minnesota Refugee Mental Health Screening Tool

Item – brief mental health tool
1. In the past month, have you felt too sad?
2. In the past month, have you been worrying or thinking too much?
3. In the past month, have thoughts about the past that kept you from doing things or spending time with others?
4. In the past month, did you have sleep problems?
5. In the past month, did you have memory problems?
If any of the above answers were yes, then ask: 6. Did any of the above stop you from doing things you need to do every day?

Background

The Minnesota Department of Health Refugee Health Program (RHP) and partners developed a brief and standardized refugee mental health screening tool that was well understood across various countries of origin resettling through the refugee program. This tool was employed as a provider-driven assessment of mental health needs in the context of the Refugee Health Assessment (RHA). The tool development followed a report from a statewide expert work group, which confirmed the utility of this approach and a gap in existing tools to address the need. Details on the development process can be accessed at Mental Health Screening: Domestic Refugee Health Guidance (www.health.state.mn.us/communities/rih/guide/mentalhealth.html).

A primary goal of the pilot was to integrate the two versions of the screening tool into a unified tool that could standardize refugee mental health screening in Minnesota. The process and findings are detailed in this report. In addition, the pilot integrated a quality improvement approach to develop best practices for a statewide mental health screening roll-out. The scope of that encompassed provider training, clinic workflow implementation, language access considerations, follow-up protocols for those with identified needs, electronic medical record (EMR) integration, and data tracking

Methods

Pilot sites and participants

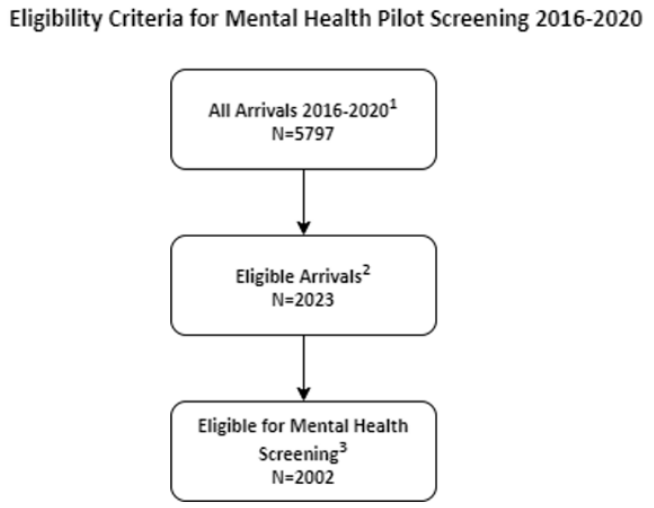
The participating pilot clinics were selected based on capacity to participate, diversity of arriving refugees, and incorporation of different health systems and workflows. The four participating clinics were located in Hennepin (1 site), Ramsey (2 sites), and Olmsted (1 site) counties. Three health systems and two public health clinics were represented; one site integrated care at a public health clinic and a health system, and one was a teaching clinic. During the pilot period, the four clinics generally conduct approximately 60% of RHA in the state.

Screening eligibility

New arrivals were eligible for the mental health screening if they received their RHA at a participating clinic and met the below criteria:

- Arrived between January 1, 2016 – January 31, 2020.
- Refugee whose first state of resident after resettlement is Minnesota (primary refugee).
- Ages 18 and older at arrival.
- No prior mental health diagnosis. Those with an existing diagnosis were connected to care through an existing referral process.
- Physically and cognitively able to answer questions.

Figure 1. Eligibility criteria for Mental Health Pilot Screening 2016-2020



- Eligibility criteria includes primary refugee arrivals and special immigrant visa holders, age 18 or older at time of arrival, being screened at participating clinics, and eligible for refugee health screening (did not move out of state, move to unknown destination, was unable to be located, never arrived, had no insurance, or died before screening).
- Exclusion criteria for mental health screening and subsequent data analysis include having a cognitive impairment, pre-existing mental health condition, moved elsewhere or was screened at a non-participating clinic.

Screening tools

The screening tool is intended to gauge the distress level of new arrivals, as distinct from a specific diagnosis or the presence of Serious and Persistent Mental Illness (SPMI)[1]. Two versions of the tool were created; four or five concrete symptom-based questions were included in Version A and B, respectively, along with one functionality item (Table 2 and 3).

Table 2. Version A screening measure items and variables names

Item	Variable
1. In the past month, have you felt very sad?	SAD-1
2. In the past month, have you been worrying or thinking too much?	ANX-2
3. In the past month, have you had any bad dreams or nightmares?	SLE-3
4. In the past month, have you avoided situations that remind you of the past?	AVD-4
5. In the past month, do any of these things make it difficult to do what you need to do on a daily basis? (Prompt: Are you able to take care of yourself and your family?)	FUN-5

Table 3. Version B screening measure items and variables names

Item	Variable
1. In the past month, did you feel sad too much?	SAD-6
2. In the past month, did you worry/think too much?	ANX-7
3. In the past month, did you have sleep problems?	SLE-8
4. If yes, in the past month, did you have dreams/nightmares?	DRE-9
5. In the past month, did you have memory problems?	MEM-10
6. In the past month, did any of the above stop you from doing things you need to do every day?	FUN-11

Administration of screening

Version A was implemented at three clinics and Version B was implemented at one clinic. The screening was verbally administered by a provider (doctor, nurse practitioner, or physician assistant) through the assistance of an interpreter. One of the intended benefits of this approach is to create space for normalization and psychoeducation within the clinical moment. Generally, the screening took place during the second of the two appointments in the RHA process. Responses for each item were coded as "yes" or "no"; and, the final screening outcome, integrating clinical judgement, was recorded in the patient's electronic health record. The item responses, screening outcome, and follow-up (or referral) status were submitted to RHP in a standardized data collection form. The form was modified to clearly capture the screening outcome as "positive" or "negative."

Interpretation of screening results

For both versions, two or more endorsed items was considered a positive screening outcome (from here on called *algorithm-determined* threshold). For Version B of the screening measures, "dreams/nightmares" (DRE-9) was not used to determine a positive screen since it relied on a specific answer to a previous question "sleep problems" (SLE-8) and therefore was not systematically asked. Because the tool was designed to be used in conversation with a provider, there was also a formal opportunity for providers to integrate clinical judgement in the screening results. Three of the four participating clinics submitted final screening outcomes, so we were only able to compare this for 902 (54%) of screened participants. The *algorithm-determined* threshold agreed with the final screening outcome for 884 (96%) participants for which final screening outcomes were submitted (Table 4).

Table 4. Algorithm vs. Provider-determined Screening Outcome

		Provider Outcome	
		Positive	**Negative**
Algorithm Outcome	**Positive**	112 (94%)	11
	Negative	7	772 (99%)

Data analysis

A descriptive analysis was conducted to estimate the prevalence of reported mental health distress among participants by age, gender, and country of origin. Cross-cultural measurement analyses using the Rasch Model was then conducted with a cleaned dataset (Total N=1,832) to accomplish the following: 1) evaluate the reliability and functionality of both brief measures (Version A, N=1,004; Version B, N=828), and 2) assess whether respondents across six distinct countries of origin responded to the items on the two brief refugee mental health screening measures as intended and consistently across groups. A total of six countries of origin (Burma, Bhutan, Ethiopia, Iraq, Somalia, and the DR Congo) were included in the analyses to construct equivalent metrics for comparing the groups to one another, which required an N = > 100. Differential item functioning (DIF) analyses was performed to assess item equivalence (e.g., rates of endorsement) across the six countries of origin on both brief measures.

Several statistics were examined to assess psychometric properties of each brief measure. Infit mean square and outfit mean square (MNSQ) compared and contrasted the fit of the observed data to the expected fit calculated by the Rasch model. Infit and outfit MNSQ fit statistics should fall between 0.5 – 1.5 to be productive for measurement (Linacre, 2005). Standardized MNSQ fit statistics (ZSTD) and point measure correlation and item discrimination were evaluated to identify misfitting and poorly differentiating items. To explore how response categories were used by persons who completed each measure, use of categories was examined to identify infrequently or improperly used response options. Average measure of item difficulty, person ability and step calibration for each response category was reviewed to these values fell within acceptable ranges.

Person statistics for all six countries of origin were generated to identify differences in average person measures, average MNSQ and ZSTD, and separation reliability measures. Separation reliability is comparable to Cronbach's alpha and ideally is >.90 but adequate if falls between 0.80 – 0.90 (Linacre, 2005). Item and person maps were created to depict how items rank from most to least difficult to endorse. These maps also indicate which symptoms persons were less likely to endorse, and the average level of each trait (e.g., depression, anxiety, etc.) related to mental health distress was reported by each country of origin. Differential item functioning (DIF) was performed to examine whether observed differences between the six countries of origin were explained by the item differences (or non-equivalence). A rule of thumb utilized to assess significant DIF was logit values of >0.4 and t statistics of >1.95 (+/-).

Item fit statistics for each brief measure are provided in Table 7 and 8 (refer to Appendix A). For the Version A screening tool, items ranged from 0.94 and 1.41 in mean infit MNSQ and .65 and 1.35 in mean outfit MNSQ; whereas, all items on the Version B measure were between 0.92 and 1.13 in mean infit MNSQ and 0.61 and 1.18 in mean outfit MNSQ. For details of the cross-cultural analysis methods, see Appendix A. All data analysis was performed utilizing R statistical software 4.0.2 (R core team, 2020) and WINSTEPS statistical software 4.6.1 (Linacre, 2020).

Results

Descriptive analysis

Between January 2016 and January 2020, 5,797 new arrivals with refugee status resettled in Minnesota, and 2,002 (34%) were eligible for the mental health pilot screening (Figure 1). A total of 1,672 (84%) eligible arrivals were screened for mental health using one of the two mental health screening tools (Table 5).

Table 5. Screening Outcome for Mental Health Screening Pilot 2016-2020

	No (% of total)
Participants	2,002
Screened	1,672 (84%)
Not Screened	330 (16%)
Reason for Not Screened	
Failed Appointment	28 (8%)
Not offered by Provider	102 (31%)
Other	106 (32%)
Unknown	94 (28%)

Although the majority (84%) of participants were successfully screened, reasons for not screening a participant include not being offered by the provider, missing appointments, other pressing health needs, and lack of time.

Table 6. Descriptive Summary of Mental Health Screening Pilot 2016-2020

	No (% of total)		
	All Participants	**Positive for Distress**	**% Positivity**
Total	1672	215	13%
Age at US Arrival (yrs)			
18-24	432 (26%)	38 (18%)	9%
25-40	757 (45%)	80 (37%)	11%
41-60	352 (21%)	67 (31%)	19%
61+	131 (8%)	30 (14%)	23%
Gender			
Male	767 (47%)	127 (59%)	14%
Female	866 (53%)	88 (11%)	11%
Country of Origin			
Afghanistan	39 (2%)	11 (5%)	28%
Bhutan	76 (5%)	11 (5%)	14%
Burma	544 (33%)	70 (33%)	13%
Congo DR	132 (8%)	25 (12%)	19%
Eritrea	52 (3%)	6 (3%)	12%
Ethiopia	155 (9%)	20 (9%)	13%
Iraq	59 (4%)	21 (10%)	36%
Somalia	484 (29%)	33 (15%)	7%
Ukraine	45 (3%)	4 (2%)	9%
Other	86 (5%)	14 (7%)	16%

Overall, 215 (13%) participants screened positive for mental health distress as defined by the algorithm-determined threshold. The highest percent positivity was among participants from Iraq and Afghanistan, with 36% and 28% screening positive for mental health distress, respectively. Although the majority of the participants were younger, with a mean age at arrival of 35.9 years, those who were 61 years or older had the highest percent positivity at 23% (Table 6).

A review of relevant mental health distress among refugees found in literature revealed prevalence broadly similar to or above the positive screening rates in this pilot. Of note, no formal meta-analysis was performed, and studies frequently measured diagnoses of anxiety, depression, and PTSD. References for this literature review are included in a secondary reference section (1-17).

Cross-cultural measurement analysis

Conducting screening with ethnically and culturally diverse populations, particularly when using translated instruments and/or interpreters for administration, poses challenges. A fundamental issue is the potential that the measures developed to evaluate a given construct (e.g., mental health distress) in one particular cultural group may not be assessing the same construct in other cultural groups. Cross-cultural measurement analysis was conducted with a cleaned dataset (Total N=1,832) to evaluate the reliability and functionality of two brief refugee mental health screening measures (Version A – N=1,004; Version B – N=828). The analyses also examined how effectively the screening measures identified mental health distress in refugees from six distinct countries of origin including Burma, Bhutan, Ethiopia, Iraq, Somalia, and the DR Congo.

Results indicated the **mean level of mental health distress was greater for refugees from Iraq, the DR Congo, and Bhutan than for refugees from Burma, Ethiopia, and Somalia on both measures**. These findings are consistent with previous literature that has found elevated rates of depression, anxiety, and PTSD reported among Iraq, Bhutan, and the DR Congo as compared to other countries of origin in the project. Although some mean differences in reported mental health distress were found across refugees from six countries of origin, **the average fit of persons was found to be appropriate for refugees overall.** Fit of persons was poorer for refugees from Bhutan and Iraq on both measures, likely a function of smaller sample sizes (<100) for these groups.

Rates of endorsement on the screening items differed by country of origin. “Worry/think too much” (ANX-2) was less endorsed for refugees from Burma and Bhutan, “sleep problems” (SLE-3) was less endorsed for refugees from Ethiopia and Somalia, and “memory problems” (MEM-10) was less endorsed by refugees from Iraq, Somalia, and the DR Congo. Refugees from Iraq were less likely to endorse FUN-5 item, and the DR Congo cultural group were less likely to endorse SAD-1 item. Overall, **item equivalence across refugee countries of origin differed on three key items: “avoidance” (AVD-4) on the Version A tool, and “dreams/nightmares” (DRE-9) and “memory” (MEM-10) on the Version B tool**. After the AVD-4 item was removed, the fit statistics and reliability of the Version A tool were reduced. Removal of the MEM-10 item on the Version B tool also resulted in poorer fit and reliability of the measure. The findings confirm the avoidance and memory items should be retained as they contributed to accurately

identifying mental health distress in refugees across countries of origin. The Version B tool did not worsen in fit statistics or reliability when DRE-9 was removed. Thus, it is recommended this item be removed from the scale or merged with the SLE-8 item (refer to Tables 11 and 12 in Appendix A).

Discussion

Psychometrically sound and cross-culturally equivalent screening instruments are essential to accurately identify the extent to which traumatic experiences and resulting mental health distress affect the well-being of refugees; and to what extent endorsement of mental health distress may differ across specific cultural groups. This measurement analyses of two brief measures developed to screen for mental health distress in recently resettled adult refugees evaluated the cross-cultural utility and equivalence across six distinct countries of origin: Burma, Bhutan, Ethiopia, Iraq, Somalia, and the DR Congo. Of note, each of these countries includes varied and intersecting cultures; for the purpose of analysis and discussion, the term "countries of origin" is used to indicate those with a shared country of origin.

Fit statistics confirmed adequate functioning of both brief measures in screening for mental health distress for newly arrived refugees. Findings revealed reliabilities for both measures were sufficient; however, the Version A screening tool demonstrated higher item and person separation reliabilities. Analysis of DIF indicated there were differences in how strongly each cultural group endorsed symptoms on both measures. Item equivalence across countries of origin was found to differ more substantially on three key items: AVD-4 on the Version A tool, and DRE-9 and MEM-10 on the Version B tool. After the AVD-4 item was removed, the fit statistics and reliability of the Version A tool was reduced. Removal of the MEM-10 item on the Version B tool also resulted in worse fit of the measure. Consequently, these findings confirm the avoidance and memory items should be retained as it contributed to accurately identifying mental health distress in newly arrived refugees. The Version B measure did not worsen in fit statistics or reliability when DRE-9 was removed. Thus, it is recommended this item be removed from the scale or merged with the SLE-8 item.

Limitations

Although results supported the utility and reliability of both measures, findings also shed light on the non-equivalence of items and differences in the endorsement of certain items across the six country of origin groups. That is, refugee countries of origin varied considerably in which screening items they endorsed underscoring the notion that mental health is understood and functions differently across cultures. DIF findings raised issues of whether mental health distress can be conceptualized similarly across cultures; and, whether it can be effectively screened for with a standardized measure.

Final recommendations

In light of the above analyses, the Minnesota Refugee Health Program will integrate and implement a consolidated screening tool in its State Refugee Mental Health Screening Guidance; the final question items are listed in Table 1. Provider feedback from the pilot phase endorsed some perceived challenges in patient understanding of the item regarding avoidance; and cultural liaisons and research partners collaborated in a rephrasing for clarify to that item.

An important limitation of this analysis is the lack of validation for the instrument. Expanded data collection through state-wide implementation of the final mental health screening instrument in parallel with validated diagnostic instruments is a suggested next step. This would allow for the confirmation of whether current screening items are valid and predictive of active mental health distress.

Appendix A

The concept of cross-cultural equivalence has several categories, including conceptual, functional, construct operationalization, item, and scalar (Hui & Triandis, 1985). Conceptual equivalence is defined as a construct has a similar meaning in different cultures; and functional equivalence must indicate similar precursors, consequents, correlates, and goals. These two equivalences are the first requirements for cross-cultural comparisons and are related to underlying theories of a measure. Evidence that a construct is operationalized in the same manner in different countries of origin and is akin to conceptual and functional equivalence.

Item and scalar equivalence involve an investigation of the psychometric properties of a measure. Every item must denote the same meaning across cultures. Item equivalence enables meaningful numerical comparisons between cultures or other groups (Hui & Triandis, 1985; Reise, Widaman & Pugh, 1993). Scalar equivalence indicates a particular score on a measure represents the same degree, intensity, or magnitude of the construct across groups. Scalar equivalence is also the most difficult to confirm but is a necessary prerequisite for diagnostic instruments.

Rasch and item response theory (IRT) models can be estimated using joint maximum likelihood (JML) techniques (Fischer & Molenaar, 1997; Wright & Masters, 1982), one can compare subgroups of respondents regardless of the raw score distributions. To that end, Rasch and IRT analysis is more appropriate technique for examining cross-cultural equivalency of psychiatric measures, which are likely to be non-normal distributions. IRT models also posit more stringent sets of measurement invariance constraints because they account for the item difficulties, which are ignored in CFA (Reise et al., 1993). An analysis of the properties of a measure for different subgroups of the sample can indicate differential fit to the Rasch model, and differential item functioning (DIF) can determine whether the item is endorsed similarly across different groups (Gerber et al., 2002). Rasch model is a log odds model that utilizes principles of inverse probability and conjoint additivity to calculate difficult of individual persons and items. The technique measures log odds of a person selecting any category on an item as an additive function of the person's ability and item difficulty of the rating scale response categories. All a whole, these methods help shed light on possible differences in how constructs are conceptualized in different cultures and whether items on a measure are equivalent.

Table 7. Item Fit Statistics by Country of Origin for Mental Health Screening Tool Version A

	Mean Infit (MNSQ)	Mean Infit (ZSTD)	Mean Outfit (MNSQ)	Mean Outfit (ZSTD)	PTMEA Correlation	Discrimination
Country of Origin						
Burma	1	0.97	1.35	1.33	0.62	1.19
Bhutan	0.94	0.73	1.12	1.45	0.47	0.98
Ethiopia	1.02	1.41	0.65	1.42	0.57	1.12
Iraq	1.06	1.37	1.23	1.56	0.54	0.94
Somalia	0.97	1.24	1.12	1.23	0.6	1.15
DR Congo	1.03	1.38	1.17	1.36	0.52	0.97
All Participants	0.99	1.14	0.97	1.18	0.59	1.09

Table 8. Item Fit Statistics by Country of Origin for Mental Health Screening Tool Version B

	Mean Infit (MNSQ)	Mean Infit (ZSTD)	Mean Outfit (MNSQ)	Mean Outfit (ZSTD)	PTMEA Correlation	Discrimination
Country of Origin						
Burma	0.99	0.85	1.03	1.41	0.6	1.1
Bhutan	0.92	0.81	1.15	1.5	0.44	0.95
Ethiopia	1.05	1.27	0.61	1.39	0.56	1.04
Iraq	1.13	1.42	1.18	1.48	0.51	0.9
Somalia	0.96	1.31	1.09	1.34	0.59	1.11
DR Congo	1.05	1.29	0.79	1.23	0.54	0.99
All Participants	0.97	1.21	1.06	1.29	0.57	1.02

Summary statistics for persons overall and for each country of origin group are provided in Tables 7 and 8. The mean level of mental health distress was greater for countries of origin from Iraq, the DR Congo, and Bhutan than mean level of mental health distress reported by persons from Burma, Ethiopia, and Somalia on both measures. This may be related in part to item and person fit of the measures accurately identifying individuals with mental health distress in these cultural groups. Additionally, these findings are consistent with previous literature that has found elevated rates of depression, anxiety and PTSD reported among Iraqis, Bhutanese and the DR Congolese as compared to other countries of origin in the project.

Table 9. Person Summary Statistics by Country of Origin for Mental Health Screening Tool Version A

	Mean Infit (ZSTD)	Mean Outfit (ZSTD)	RMSE	Separation	Reliability
Country of Origin					
Burma	1.02	1.31	0.023	3.22	0.97
Bhutan	1.08	1.36	0.027	2.57	0.87
Ethiopia	0.99	1.04	0.031	3.16	0.93
Iraq	1.12	1.41	0.034	2.71	0.89
Somalia	0.96	1.11	0.021	2.95	0.91
DR Congo	1.05	1.17	0.028	3.03	0.92
All Participants	1.01	1.19	0.026	3.11	0.93

Table 10. Person Summary Statistics by Country of Origin for Mental Health Screening Tool Version B

	Mean Infit (ZSTD)	Mean Outfit (ZSTD)	RMSE	Separation	Reliability
Country of Origin					
Burma	1.06	1.28	0.025	3.17	0.95
Bhutan	1.12	1.43	0.036	2.52	0.86
Ethiopia	1.01	1.14	0.033	3.01	0.92
Iraq	0.77	1.31	0.04	2.49	0.86
Somalia	0.85	1.22	0.029	2.96	0.91
DR Congo	1.04	1.27	0.026	2.83	0.89
All Participants	0.98	1.26	0.032	2.89	0.9

Average fit of persons was found to be appropriate for persons overall and for countries of origin from Burma, Ethiopia, the DR Congo, and Somalia on both brief screening tools. Fit of persons was poorer for those from Bhutan and Iraq on both measures. This result was likely a function of smaller sample sizes (<100) for these groups. Results indicated the Burmese had the highest separation reliability (3.22, 0.97) on the Version A screening tool followed by countries

of origin from Ethiopia (3.16, 0.93), the DR Congo (3.03, 0.92), Somalia (2.95, 0.91), Iraq (2.71, 0.89), and Bhutan (2.57, 0.87). For the Version B screening tool, separation reliability was greatest for those from Burma (3.17, 0.95), Ethiopia (3.01, 0.92), Somalia (2.96, 0.91), the DR Congo (2.83, 0.89), Bhutan (2.52, 0.86) and Iraq (2.49, 0.86). Average infit and outfit for persons on both measures were found to be adequate, ranging between 1.01 and 1.19 for the Version A measure and 0.98 to 1.26 for the Version B tool. The fit statistics suggested the Version A screening tool performed slightly better than the Version B measure, however; there was not a substantial difference in fit found between the two brief screening measures. Findings of the DIF analyses including item difficulty and corresponding standard error for each item are presented in Table 11 for the Version A measure and Table 12 for the Version B measure. Items with highly significant DIF are indicated with an asterisk in each table. The DIF contrast is the difference between the item difficult measures for each cultural group, and the statistical significance of the contrast is revealed in the t-statistic and corresponding p-value of the item.

On the Version A screening measure, item difficulty substantially varied across cultural groups. A common pattern found was that AVD-4 was difficult for all countries of origin to endorse (on average), which indicated the item was less likely to be endorsed across all persons. ANX-2 was more difficult for the Burmese and Bhutanese; whereas, SLE-3 was more difficult to endorse for refugees from Ethiopia and Somalia. Iraqis were less likely to endorse FUN-5 item, and refugees from the DR Congo were less likely to endorse SAD-1 item.

Results for the Version B screening measure revealed a similar pattern overall. All countries of origin were less likely to endorse DRE-9, which showed high DIF among all groups. MEM-10 was difficult for refugees for all countries of origin to endorse except Ethiopia. Additionally, there were differences in DIF for specific cultural groups. Refugees from Burma and Bhutan were less likely to endorse ANX-7; Somalis were less likely to endorse SLE-8; Iraqis were less likely to endorse FUN-11; and the DR Congolese were less likely to endorse SAD-6.

Average measure of item and person difficulty was found to increase with each response category indicating expected category use across cultural groups. For the Version A measure, the step calibration for the avoidance item (AVD-4) differed in endorsement between cultural groups. The Somali, Iraqi, and Burmese refugees endorsed this item less than Bhutanese and the Congolese. However, when the item was removed and analyses re-run, the measure showed worse fit statistics indicating this item is an important aspect to retain despite the irregular endorsement across cultural groups.

The item assessing memory problems (MEM-10) on the Version B measure were endorsed less by Iraqis, Somalis, and the Congolese as compared to the Burmese, Bhutanese, and Ethiopians. Similar to the AVD-4 item, when MEM-10 was removed the SET B measure had poorer fit overall, thus it is recommended this item be retained as it appears to be a necessary area for detecting mental health distress. Additionally, the bad dreams/nightmares item (DRE-9) on the SET B measure endorsed less by Somalis, the Congolese, and Iraqis as compared to the Burmese, Bhutanese, and Ethiopians. When the item was removed, the fit of the measure was unaffected. As a result, it is suggested this item (DRE-9) could be removed or merged with the sleep item (SLE-9) with little to no negative impact in the screening tool's ability to detect for mental health distress.

Table 11. Item DIF by Cultural Group for Mental Health Screening Tool Version A

	Item Name	Item DIF	SE	DIF Contrast	Total SE	t	p-value
Country of Origin							
Burma	SAD-1	2.49	0.36	1.44	0.59	4.56	<0.001
Burma	ANX-2*	3.27	0.56	1.62	0.84	5.34	<0.001
Burma	SLE-3	2.89	0.45	1.38	0.72	3.39	0.005
Burma	AVD-4*	3.34	0.63	1.97	1.04	5.72	<0.001
Burma	FUN-5	2.26	0.17	1.14	0.42	3.03	0.002
Bhutan	SAD-1	2.81	0.44	-1.21	0.72	-2.95	0.025
Bhutan	ANX-2*	3.13	0.52	1.49	0.93	3.22	0.011
Bhutan	SLE-3	2.47	0.28	-1.32	0.67	2.87	0.032
Bhutan	AVD-4*	3.26	0.58	1.87	0.98	5.49	<0.001
Bhutan	FUN-5	2.39	0.24	-1.29	0.65	1.65	0.063
Ethiopia	SAD-1	3.05	0.47	1.71	0.79	4.55	<0.001
Ethiopia	ANX-2	2.91	0.46	1.63	0.7	4.18	0.002
Ethiopia	SLE-3*	3.12	0.52	1.85	0.91	4.69	<0.001
Ethiopia	AVD-4*	3.15	0.52	1.81	0.85	4.63	<0.001
Ethiopia	FUN-5	2.2	0.16	-1.07	0.39	1.74	0.061
Iraq	SAD-1	2.93	0.46	-1.42	0.76	2.2	0.017
Iraq	ANX-2	3.09	0.51	-1.81	0.87	2.37	0.011
Iraq	SLE-3	2.77	0.43	-1.53	0.8	-2.82	0.009
Iraq	AVD-4*	3.56	0.72	2.06	1.44	4.15	<0.001
Iraq	FUN-5*	3.29	0.57	1.66	1.16	5.24	<0.001
Somalia	SAD-1	2.66	0.4	-1.46	0.69	-2.98	0.013
Somalia	ANX-2	2.72	0.42	1.55	0.73	3.13	0.008
Somalia	SLE-3*	3.18	0.47	1.87	0.92	3.87	0.001
Somalia	AVD-4*	3.96	0.84	-2.94	2.03	-5.65	<0.001
Somalia	FUN-5	2.79	0.43	-1.57	0.76	-3.28	0.005
Congo DR	SAD-1*	3.16	0.52	1.9	0.98	4.77	<0.001
Congo DR	ANX-2	2.98	0.46	1.81	0.85	3.35	0.004
Congo DR	SLE-3	2.57	0.39	1.48	0.61	3.17	0.01
Congo DR	AVD-4*	3.41	0.66	2.35	1.57	5.46	<0.001
Congo DR	FUN-5	3.02	0.34	-1.66	0.82	-4.73	<0.001

Table 11 shows the item DIF by cultural group for Version A of the mental health screening tool.
* Indicates key finding

Table 12. Item DIF by Cultural Group for Mental Health Screening Tool Version B

	Item Name	Item DIF	SE	DIF Contrast	Total SE	t	p-value
Country of Origin							
Burma	SAD-6	2.34	0.39	1.51	0.65	4.71	<0.001
Burma	ANX-7*	3.21	0.54	1.58	0.8	5.27	<0.001
Burma	SLE-8	2.91	0.46	-1.45	0.77	-3.42	0.003
Burma	DRE-9*	3.26	0.55	1.92	1.01	5.66	<0.001
Burma	MEM-10*	3.11	0.49	-1.82	0.78	-4.54	<0.001
Burma	FUN-11	2.23	0.14	1.09	0.35	2.92	0.005
Bhutan	SAD-6	2.89	0.46	-1.27	0.78	-3.07	0.012
Bhutan	ANX-7*	3.17	0.62	1.56	1.03	3.52	0.004
Bhutan	SLE-8	2.45	0.29	-1.38	0.74	2.99	0.015
Bhutan	DRE-9*	3.21	0.59	1.91	1.01	5.57	<0.001
Bhutan	MEM-10*	3.23	0.6	-1.96	1.12	-5.71	<0.001
Bhutan	FUN-11	2.41	0.27	1.23	0.69	1.83	0.058
Ethiopia	SAD-6	2.41	0.34	1.48	0.54	4.03	<0.001
Ethiopia	ANX-7	2.78	0.39	1.53	0.65	4.11	<0.001
Ethiopia	SLE-8	2.84	0.45	-1.62	0.71	-3.92	0.001
Ethiopia	DRE-9*	3.1	0.56	1.74	0.77	4.34	<0.001
Ethiopia	MEM-10	3.02	0.48	1.82	0.91	4.19	<0.001
Ethiopia	FUN-11	2.24	0.25	1	0.41	2.87	0.019
Iraq	SAD-6	2.87	0.42	-1.34	0.61	-3.06	0.013
Iraq	ANX-7	2.74	0.46	1.72	0.59	3.03	0.013
Iraq	SLE-8	2.89	0.5	-1.69	0.92	-2.95	0.015
Iraq	DRE-9*	3.29	0.67	-1.37	1.23	-4.02	<0.001
Iraq	MEM-10*	3.36	0.65	1.85	1.62	4.08	<0.001
Iraq	FUN-11*	3.19	0.55	1.57	1.13	5.06	<0.001
Somalia	SAD-6	2.52	0.37	1.43	0.54	2.87	0.022
Somalia	ANX-7	2.66	0.4	-1.49	0.63	-2.94	0.019
Somalia	SLE-8*	3.11	0.61	1.93	1.01	3.98	<0.001
Somalia	DRE-9*	3.75	0.72	-2.81	1.9	-5.48	<0.001
Somalia	MEM-10*	3.32	0.57	2.61	1.74	4.33	<0.001
Somalia	FUN-11	2.84	0.48	-1.5	0.62	-3.05	0.009
Congo DR	SAD-6*	3.18	0.56	1.82	0.81	4.64	<0.001
Congo DR	ANX-7	2.53	0.36	-1.55	0.63	-2.88	0.024
Congo DR	SLE-8	2.45	0.32	-1.42	0.54	-2.81	0.025
Congo DR	DRE-9*	3.61	0.74	2.48	1.77	5.58	<0.001
Congo DR	MEM-10*	3.27	0.54	-2.27	1.45	-4.97	<0.001
Congo DR	FUN-5	2.85	0.46	1.62	0.79	3.46	0.004

Table 12 shows the item DIF by cultural group for Version B of the mental health screening tool.

* Indicates key finding

References

Fischer G. H., Molenaar I. W. (1997). *Rasch Models: Foundations, recent developments, and applications.* New York: Springer Verlag.

Gerber B., Smith E. V., Girotti M., Pelaez L., Lawless K., et al. (2002). Using Rasch measurement to investigate the cross-form equivalence and clinical utility of Spanish and English versions of a Diabetes Questionnaire: A pilot study. *Journal of Applied Measurement,* 3, 243–271.

Hui CH, Triandis HC. (1985). Measurement in cross-cultural psychology: a review and comparison of strategies. *Journal of Cross Cultural Psychology, 16,* 131–152

Linacre J. M. (2020). WINSTEPS Rasch measurement statistical software computer program 4.6.1.

R Core Team (2020). R statistical software 4.0.2 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.r-project.org/

Reise, S. P., Widaman, K. F., Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychol Bull 114(3), 552-566.*

Wright B. D., Masters G. N. (1982). *Rating scale analysis.* Chicago: MESA Press.

Secondary References (Literature Overview of SPMI Rates)

1. Slewa-Younan, S., Yaser, A., Guajardo, M.G.U. et al. The mental health and help-seeking behaviour of resettled Afghan refugees in Australia. Int J Ment Health Syst 11, 49 (2017) doi:10.1186/s13033-017-0157-z
2. Malekzai, A.S.B., Niazi, J.M., Paige, S.R. et al. J Trauma Stress (1996) 9: 891. https://doi-org.ezp2.lib.umn.edu/10.1007/BF02104111
3. Gernaat, Hajo B.P.E. et al. "[Many psychiatric disorders in Afghan refugees with residential status in Drenthe, especially depressive disorder and post-traumatic stress disorder]." Nederlands tijdschrift voor geneeskunde 146 24 (2002): 1127-31.
4. Mghir, R., Freed, W., Raskin, A., & Katon, W. (1995). Depression and Posttraumatic Stress Disorder Among a Community Sample of Adolescent and Young Adult Afghan Refugees. The Journal of Nervous and Mental Disease, 183(1), 24-30.
5. Gerritsen, A.A.M., Bramsen, I., Devillé, W. et al. Soc Psychiat Epidemiol (2006) 41: 18. https://doi-org.ezp1.lib.umn.edu/10.1007/s00127-005-0003-5
6. Liebman, S. An Overview of the Unique Mental Health Challenges facing the Bhutanese Refugee Population. November 2014. https://med.virginia.edu/family-medicine/wp-content/uploads/sites/285/2017/02/Spencer-Liebman-Mental-Health-Bhutanese_Web.pdf
7. Ao, T., Shetty, S., Sivilli, T., Blanton, C., Ellis, H., Geltman, P., . . . Lopes Cardozo, E. (2016). Suicidal Ideation and Mental Health of Bhutanese Refugees in the United States. Journal of Immigrant and Minority Health, 18(4), 828-835.
8. K Allden, C Poole, S Chantavanich, K Ohmar, N N Aung, and R F MollicaHarvard Program in Refugee Trauma, Cambridge, Mass 02138, USA. "Burmese political dissidents in Thailand: trauma and

survival among young adults in exile.", American Journal of Public Health 86, no. 11 (November 1, 1996): pp. 1561-1569.

9. Kim, I. (2018). Behavioral health symptoms among refugees from Burma: Examination of sociodemographic and migration-related factors. Asian American Journal of Psychology, 9(3), 179-189. http://dx.doi.org.ezp3.lib.umn.edu/10.1037/aap0000103
10. Schweitzer, R., Brough, M., Vromans, L., & Asic-Kobe, M. (2011). Mental Health of Newly Arrived Burmese Refugees in Australia: Contributions of Pre-Migration and Post-Migration Experience. Australian and New Zealand Journal of Psychiatry, 45(4), 299-307.
11. Johnson K et al. Association of Sexual Violence and Human Rights Violations with Physical and Mental Health in Territories of the Eastern Democratic Republic of the Congo. Journal of the American Medical Association 2010; Vol. 304. No. 5: 553-62. Available at: https://jamanetwork.com/journals/jama/fullarticle/186342
12. Greene, C. Ventevogel, P. Tol, W. Mental Health and Psychosocial Wellbeing in Congolese Refugee Survivors of GenderBased Violence: A Desk Review. October 2016. https://pdfs.semanticscholar.org/cf8d/b3c6b7e6cca5661b7f67bfd558e6ad15b014.pdf
13. Chernet, A., Pfeiffer, C., Probst-Hensch, N., & Labhardt, A. (2016). Mental health and resilience among newly arrived Eritrean refugees in Switzerland. European Journal Of Public Health, 26(S1), 400-401.
14. Kasturi, P. Iyengar, R. Haile, A. Mental Health Issues and Substance Abuse in the Ethiopian Community Diaspora: Economic Costs, Public Health Perspective and Policy. December 2014. American Research Insitute for Policy Development. https://pdfs.semanticscholar.org/9849/9706561d5a7eb6ebb6f535d817b8206b21e7.pdf
15. Alhasnawi, Salih, et al. The prevalence and correlates of DSM-IV disorders in the Iraq Mental Health Survey (IMHS). 2, 2009, World Psychiatry, Vol. 8, pp. 97-109.
16. Taylor, E., Yanni, M., Pezzi, E., Guterbock, A., Rothney, C., Harton, M., . . . Burke, J. (2014). Physical and Mental Health Status of Iraqi Refugees Resettled in the United States. Journal of Immigrant and Minority Health, 16(6), 1130-1137.
17. Korostiy V., Zavorotniy V., Polishuk V. Pharmacotherapy in complex treatment and rehabilitation of PTSD. INTERNATIONAL JOURNAL OF NEUROLOGICAL 6 (76) 2015. P. 59-71