A CLINICIAN'S COMPENDIUM OF ASSESSMENT TOOLS FOR MENTAL HEALTH CLIENTS FROM CULTURALLY AND LINGUISTICALLY DIVERSE (CALD) BACKGROUNDS

D Ryder[1], J Little[1], B Wright[2], A Pearman[3], B Willett[2]

[1]Centre for Public Health
Edith Cowan University

[2]West Australian Transcultural Mental Health Centre

[3]ASeTTS
(Association for Services to Torture and Trauma Survivors)

A project funded by the Office of Mental Health
Of the Department of Health, Western Australia

EDITH COWAN UNIVERSITY

ASeTTS
Assisting Torture and Trauma Survivors

WEST AUSTRALIAN TRANSCULTURAL MENTAL HEALTH

CONTENTS

INTRODUCTION 4

USING THE COMPENDIUM 5

CLINICAL ASSESSMENT INSTRUMENTS FOR USE WITH CALD POPULATIONS 6

GLOSSARY OF PSYCHOMETRIC TERMINOLOGY USED IN THE COMPENDIUM 7

HARVARD TRAUMA QUESTIONNAIRE (HTQ) 9

Harvard Trauma Questionnaire : Khmer, Laotian, Vietnamese 11

Hospital Anxiety and Depression Scale (HADS) 13

Hospital Anxiety and Depression Scale (HADS): Cantonese 15

Hospital Anxiety and Depression Scale (HADS): Italian 16

Hospital Anxiety and Depression Scale (HADS): Arabic 17

Hospital Anxiety and Depression Scale (HADS): Spanish 18

ARABIC OBSESSIVE-COMPULSIVE SCALE (AOCS) 19

BECK DEPRESSION INVENTORY II (BDI-II) 22

Beck Depression Inventory II (BDI-II): Arabic 24

Beck Depression Inventory (BDI): Chinese 25

Beck Depression Inventory II (BDI-II): Spanish 27

Chinese Health Questionnaire (CHQ) 29

GENERAL HEALTH QUESTIONNAIRE (GHQ) 32

General Health Questionnaire (GHQ): Arabic 34

General Health Questionnaire-30 (GHQ-30): Italian 36

General Health Questionnaire-12 (GHQ-12): Italian 38

General Health Questionnaire (GHQ): Chinese Dialects 41

EDINBURGH POSTNATAL DEPRESSION SCALE (EPDS) 43

Edinburgh Postnatal Depression Scale (EPDS): Cantonese 45

Edinburgh Postnatal Depression Scale (EPDS): Arabic 47

Edinburgh Postnatal Depression Scale (EPDS): Italian 48

APPENDIX A: 49

LIST OF EVALUATED INSTRUMENTS BY LANGUAGE

INTRODUCTION

Australia is a multicultural country, with 23% of its population born overseas and many others born in Australia who identify with different ethnic groups. An issue of concern in the mental health field, amongst others, is that services and the clinical instruments used by service providers may not be culturally appropriate for these populations, in particular where they are from culturally and linguistically diverse (CALD) backgrounds. This, it is suggested, can lead to a diagnosis of psychopathology when it is not present, failure to diagnosis psychopathology when it is present and errors in assessing severity of psychopathology (Dana, 1995). The development of appropriate tools to be used with CALD populations does not involve simply the translation of English language versions of the instrument into another language, requiring specific translations methods be applied. The instrument also needs to be assessed for reliability, validity and other psychometric properties.

The appropriate development of mental health instruments for use with CALD populations requires therefore comprehensive and often expensive processes, beyond the resources available to working clinicians. The Office of Mental Health (OMH) of the Department of Health of Western Australia (DoH, WA) funded a project that involved identifying and reviewing translated mental health instruments, evaluating as far as possible their appropriateness for use with CALD populations in Western Australia and making this information available to clinicians working within the State's mental health sector in a compendium. This is the compendium that resulted from the project.

USING THE COMPENDIUM

The compendium is presented as follows.

Following the introduction is a glossary of terms used in the compendium. After this, those instruments that are recommended for use with caution are summarised, with key references for each instrument referenced after each entry. Unless otherwise noted, it is suggested that the material be used with caution because it was developed for use in circumstances quite different to those pertaining to CALD populations in Western Australia. The suggestion that the material be used with caution should not be read as imputing any criticism of the material if they are used in the circumstances for which they were developed. For each instrument there are two sections that you are encouraged to read. The first section is titled 'background. This provides information on the instruments availability, from whom it is available, describes the instrument and how it is administered, the way it is scored, how long it takes to administer and what restrictions if any there are on using it. Following this there is a section specific to the instrument in each of the languages that it's use is deemed appropriate WITH CAUTION. Appendix A at the end of the compendium is designed in a 'user friendly' manner, which can be used by clinicians to identify instruments by language.

This compendium is the result of a detailed review of available literature, much of the technical detail having been excluded. For those interested in further detail, the full report of the project is available at:

www.mmha.org.au/watmhc/

Or by e-mail from:

David.Ryder@ecu.edu.au

Or

Bernadette.Wright@health.wa.gov.au

CLINICAL ASSESSMENT INSTRUMENTS FOR USE WITH CALD POPULATIONS

In this section, the mental health assessment instruments reviewed as part of the funded project and which can be used **WITH CAUTION** are presented. Information on their availability and accessibility and a summary of the extent to which they adequately fulfil the requirements of the benchmark criteria specified in the project final report against which they have been assessed is presented. The format in which the information is presented is described below:

NAME OF ASSESSMENT INSTRUMENT

Source Publication:
Source publication of original instrument

Source language:
Language in which original instrument was developed

Translations:

- LANGUAGES OTHER THAN ENGLISH (LOTE) INTO WHICH THE INSTRUMENT HAS BEEN TRANSLATED AND ARE INCLUDED IN THE COMPENDIUM.
- *Instrument in LOTE for which insufficient published information available and accessible.*
  (Project Management Team makes no further comment. Use of instrument not recommended).

Contact:
Where instrument may be obtained.

Copyright:
Who holds copyright.

Description:
Describes purpose and structure/format of instrument

Population :
Population groups for which translated versions have been developed

Administration:

- Information concerning form in which instrument is administered
- Scoring information
- Acceptance of instrument by identified population group
- Time required for administration
- Restrictions (if any).

Psychometric properties:
Information on scientific quality of the instrument as assessed by set of criteria developed by the Project Team and is comprehensively set out in the full Report to this Compendium.

Overall Comment:
Commentary on suitability of instrument for use with target population in context of information presented.

GLOSSARY OF PSYCHOMETRIC TERMINOLOGY USED IN THE COMPENDIUM

Concurrent validity:

The extent to which the instrument corresponds to another accepted instrument or procedure that measures the same construct. (A commonly used 'gold standard' is that of diagnosis derived from clinical interview).

Construct validity:

The relationship between the measure and some other outcome eg. those who score higher on a depression score can be predicted to have a higher number of in-patient admissions. If the measure has good construct validity than those who have higher scores on the measure will have more in-patient admissions and vice-versa.

Convergent validity:

The extent to which two measures, or two ways of measuring the same construct, agree or converge. For example, two different rating scales for depression should have similar profiles of scores for groups of people. Those scoring high for depression on one scale should also score high on the other scale.

Diagnostic validity - Area under the Curve (AUC):

AUC represents the overall diagnostic accuracy of an instrument across all cut-off scores for that instrument.

Discriminant validity:

The extent to which an instrument discriminates between disorders. For example, an instrument that purports to measure depression should be able to discriminate between depression and other psychiatric disorders.

Divergent validity:

The reverse of convergent validity. In this case, measures of two dissimilar constructs should have different outcomes. For example, high scores on a depression rating scale should be mirrored by low scores on a general life satisfaction scale.

Face validity:

Refers to whether the items in the instrument seem to relate to the constructs being assessed. That is, are the items transparent. If the meaning of items are transparent it is expected people will be less suspicious of their content and are more likely to respond.

Factor Analysis:

A statistical procedure that looks at the way in which items in an instrument 'clump' together based on the extent to which they correlate. For example, if an instrument aims to measure and discriminate anxiety and depression, then factor analysis should lead to a two factor solution, one factor derived from a close correlation between all items that measure anxiety and a second factor derived from close correlation of items that measure depression.

Internal consistency:

The extent to which scale scores are free of random error. Therefore, questions within an instrument that assess one construct should correlate strongly with other questions that measure the same construct.

Inter-rater reliability:

The instrument is administered by two or more individuals to establish the degree of consensus by those who administer it.

Predictive Validity:

The extent to which the instrument predicts outcomes over time. For example, if an instrument purports to measure severity of depression, then those individuals diagnosed as more severely depressed should show more need for medication, require more counselling etc, than those who are less severely depressed.

Sensitivity (of an instrument):

The extent to which an instrument identifies true positive cases - ie. confirmation that a client to whom the instrument is administered, has the disorder.

Specificity:

The extent to which it identifies those who are true negative cases - ie. Confirmation that a client to whom the instrument is administered does not have the disorder.

Test-retest reliability (also known as score-rescore reliability):

A measure of the stability of an instrument's scores over time, calculated by administering the test to the same person on two occasions and assessing the correlation of the two scores.

Validity:

The extent to which the instrument measures what it aims to measure.

Harvard Trauma Questionnaire (HTQ)

Source Publication:
Mollica, R. F., Caspi-Yavin, Y., Bollini, P., Truong, T., et al. (1992). The Harvard Trauma Questionnaire: Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. *Journal of Nervous & Mental Disease,180*, 111-116.

Source Language:
English

Translations:
KHMER, LAOTIAN, VIETNAMESE

Bosnian, Cambodian, Croatian, Japanese
Other languages are also under development (see below).

Contact:
Harvard Program in Refugee Trauma. Translated versions of the HTQ (and their respective manuals) can be down loaded free of charge from **http://www.hprt-cambridge.org/index.html** although free registration is required. English versions accompany the HTQ and HTQ manuals.

Arabic, Farsi, Serbo-Croatian and Russian adaptations of the HTQ have been developed by the De Vonk Psychiatric Clinic, http://www.centrum45.nl/ukindvon.php Westeinde 94, 2211XS Noordwijkerhout, The Netherlands. E-mail: mail.devonk@centrum45.nl

Copyright:
Harvard Program in Refugee Trauma

DESCRIPTION:
"The HTQ is a cross-cultural instrument designed for the assessment of trauma and torture…. symptoms associated with the diagnostic criteria for post-traumatic stress disorder (PTSD) as defined by the…DSM-III-R" (Mollica et al, 2003, p. 112).

The HTQ consists of four parts. Part I assesses 17 traumatic life events ranging from "lack of food or water" to "rape" and "torture". These events are rated on a 4-point scale: "Experienced," "Witnessed," "Heard about it," or "No." Part II asks the respondent to describe the most traumatic experience or event they have experienced during their refugee experience. Part III assesses likelihood of head injury. Part IV includes 30 trauma symptoms rated on a 4-point likert response scale anchored from 1 (*Not at all*) to 4 (*Extremely*). Seven culturally specific items have been added to the Japanese version. In the Bosnian and Croatian Veterans versions, Part IV includes the 16 DSM-IV PTSD questions and 24 additional symptom items that focus on the impact of trauma on the ability to function in everyday life.

POPULATION FOR WHICH INSTRUMENT WAS ORIGINALLY DEVELOPED:
Adult refugees and Bosnian and Croatian war veterans.

ADMINISTRATION:

- **Method:** Clinical interview, which includes the use of a self-report checklist.

- **Scoring:** Parts I, II, and III are not designed for scoring, rather they are intended to provide a means of summarising traumatic events. The total score for Part IV is the mean score for items answered with higher scores representing a greater likelihood that symptoms are associated with trauma.

- **Acceptance:** No statistical evidence. Mollica et al (1992) report that that the instrument is well accepted but people are not comfortable with disclosing information on sexual abuse.

- **Time required:** 50 – 60 minutes.

- **Restrictions:** None, although considerable clinical expertise in PTSD is strongly recommended. As a minimum Mollica et al (1992) recommend the HTQ should be administered by para-professionals under the supervision and support of psychiatrists, medical doctors, and/or psychiatric nurses.

Harvard Trauma Questionnaire:
Khmer, Laotian & Vietnamese

POPULATION ON WHICH THIS VERSION WAS TRIALED:
New and existing outpatients within a Massachusetts psychiatric clinic (administered in three Indochinese language versions: Laotian, Khmer and Vietnamese); and Vietnamese former political prisoners living within Boston.

PSYCHOMETRIC PROPERTIES

- ***Construct validity:*** Insufficient evidence. Further evidence is needed.

- ***Diagnostic validity****:* The 16 DSM-III-R trauma symptom scores showed average sensitivity and specificity when assessed against the DSM-III-R diagnosis of PTSD as determined by psychiatric staff. In a community-based sample of Vietnamese ex-prisoners of war the first 16 symptoms relating to DSM-III criteria showed perfect sensitivity and specificity and complete diagnostic accuracy (AUC) in the assessment of PTSD when assessed against the Structured Clinical Interview.

- ***Internal consistency and Inter-rater reliability:*** Estimates of internal consistency and inter-rater reliability are excellent for trauma event scores (Part I) and trauma symptom scores (Part IV) when assessed using responses pooled from the three language groups (Loatian, Khmer and Vietnamese).

- ***Test-retest reliability:*** Good for trauma event scores and excellent for trauma symptom scores. Further reliability evidence is needed.

- ***Norms:*** None.

- ***Responsiveness to change:*** Unknown.

SPECIAL NOTE:
Limited but promising evidence of internal consistency from psychiatric outpatients (Netherlands). Internal consistencies were calculated for scores from the five language versions of the trauma symptom scale. Estimates of internal consistency were good for scores from Arabic, English, Farsi, and Serbo-Croatian respondents. Internal consistency of scale scores associated with the Russian translation was poor.

OVERALL COMMENT:
The HTQ shows promise although no full validation study has been undertaken and sample sizes used have been very small. Use is recommended with caution. For an explanation of this cautionary note, see page 5, paragraph 1.

Bibliography

Fawzi, M. C. S., Murphy, E., Pham, T., Lin, L., Poole, C., & Mollica, R. F. (1997)
The validity of screening for post-traumatic stress disorder and major Depression among Vietnamese former political prisoners. Acta Psychiatrica Scandinavica, 95, 87-93.

Fawazi, M. C., Pham, T., Lin, L. Nguyen, T. V., Ngo, D., Murphy, E., & Mollica, R.F. (1997).
The validity of post-traumatic stress disorder among Vietnamese refugees. Journal of Traumatic Stress, 10, 101-108.

Hollifield, M., Warner, T. D., Lian, N., Krakow, B., Jenkins, J. H., Kesler, J., Stevenson, J., & Westermeyer, J. (2002).
Measuring Trauma and Health Status in Refugees: A Critical Review. JAMA, 288, 611 – 621.

Kleijn, W. C., Hovens, J. E., & Rodenburg, J. J. (2001).
Posttraumatic stress symptoms in refugees: assessments with the Harvard Trauma Questionnaire and the Hopkins Symptom Checklist-25 in different languages. Psychological Reports, 88, 527-32.

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. Journal of Nervous & Mental Disease, 183, 24-30.

Mollica, R. F., Caspi-Yavin, Y., Bollini, P., Truong, T., et al. (1992).
The Harvard Trauma Questionnaire: Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. Journal of Nervous & Mental Disease,180, 111-116.

Mollica, R. F., Caspi-Yavin, Lavelle, J., Tor, S., Yang, T., Chan. S., Pham, T.,Ryan, A. & de Marneffe, D. (2003).
The Harvard Trauma Questionnaire (HTQ) Manual for use withCambodian, Lao and Vietnamese versions.

Mollica, R. F., McInnes, K., Sarajlic, N., Lavelle, J., Sarajlic, I., Massagli, M. P.(1999).
Disability associated with psychiatric comorbidity and health status in Bosnian refugees living in Croatia. Journal of the American Medical Association, 282, 433-439.

Hospital Anxiety and Depression Scale (HADS)

Source Publication:
Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*, 361-370.

Source Language:
English

Translations:
ARABIC, CANTONESE, ITALIAN, SPANISH

Dutch, English, French, German, Gujerati, Japanese, Polish, Punjabi, Spanish, Swedish, and Urdu.
Use of these translations is via a licence from the publishers and the price is dependent on the volume of usage. Note: A full copy of the Swedish version can be retrieved from Salek (1998).

Contact:
NFER-Nelson Publishing Co., Ltd., Darville House, 2 Oxford Road East, Windsor, Berkshire, SL4 1DF, England.
Home Page: www.nfer-nelson.co.uk
International E-mail Enquires: information@nfer-nelson.co.uk

Copyright:
RP Snaith and Zigmond, 1983, 1992, 1994.

DESCRIPTION:
Designed to detect and distinguish between anxiety and depression and measures the severity of emotional disorder in patients with physical illness and therefore does not include somatic indicators of psychological distress.

POPULATION FOR WHICH INSTRUMENT WAS ORIGINALLY DEVELOPED:
Adults (Age range 18-65)

ADMINISTRATION:

- **Method:** Self-rating instrument.
- **Scoring:** A total of 14 items, scored on a 4-point likert scale anchored

from 0 to 3 with varying response alternatives for each item. Higher scores reflect greater emotional distress. Scoring requires 1 to 2 minutes.

- **Acceptance:** Generally well accepted, often with very high response rates.
- **Time Required:** Between 2 – 6 minutes to complete.
- **Restrictions:** Purchase restricted to Clinical Psychologists and Speech and Language Therapists, although the full English version is provided in the source publication.

Bibliography

Herrmann, C. (1997). International experiences with the hospital anxiety and depression scale: A review of validation data and clinical results. *Journal of Psychosomatic Research, 42*, 17-41.