Measuring Mental Health Literacy – A Review of Scale-Based Measures

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Journal of Mental Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>CJMH-2013-0116</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Review</td>
</tr>
<tr>
<td>Subject Area:</td>
<td>Mental Health</td>
</tr>
<tr>
<td>Further Detail:</td>
<td>Mental Health Literacy</td>
</tr>
</tbody>
</table>
Measuring Mental Health Literacy – A Review of Scale-Based Measures

Abstract

Background: Mental Health Literacy (MHL) has become an important concept in the literature; however, quantitative research on MHL methodology has been limited.

Aims: This review identified peer-reviewed papers investigating MHL, assessed psychometric attributes of scale-based measures of MHL, and the extent that studies assessed the attributes that define MHL.

Method: Two searches were conducted. First, a frequency count of publications on MHL was generated. Second, a comprehensive review of the literature to identify measures which generate a total MHL or subscale score was also conducted. A total of 13 studies were identified and their psychometric attributes determined using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN). The extent to which each measure assessed the attributes of MHL was also determined.

Results: Results showed an increase in scholarly activity regarding MHL in the past eight years. Most of the included studies failed to reported detailed information about the sample, measure development and testing to demonstrate the psychometric properties of their tool.

Conclusion: There are substantial limitations in current ability to measure MHL and there is significant scope for the development and evaluation of psychometrically robust measures that assess the relevant attributes of MHL.

Declaration of interests: No conflicts of interests or acknowledgements.
Keywords

mental health, mental health literacy, measure, psychometric
Mental Health Literacy (MHL) is a multifaceted concept which broadly refers to an individual’s knowledge and beliefs about mental health that aid recognition, management or prevention (Jorm et al., 1997). MHL provides a useful structure for understanding factors that may affect an individual’s mental health and help seeking behaviour. Globally, a high prevalence of mental illness exists, which significantly impacts on quality of life for individuals and communities (World Health Organization, 2006, 2008). However, compared with physical health, the general population has limited knowledge of the factors that affect mental health, which may compromise their ability to recognise disorders and seek effective treatments (Bartlett et al., 2006; Farrer et al., 2008; Jorm et al., 2005).

The original measure of MHL was developed by Jorm et al. (1997) and since that time, a number of other measures have been developed. To date, however, there has been no formal examination of the psychometric properties of these measures. Reliable and valid measurement of MHL is a key step in the promotion of mental health at individual and community levels; with accurate measurement providing insight into the effects of mental health education initiatives and capacity for meaningful and readily interpretable comparisons between individuals. The aim of this review is to assess the psychometric properties, scope and overall quality of current scale-based measures of MHL.

Jorm, et al. (1997) proposed the concept of MHL as an extension of the term Health Literacy (HL), which refers to an individual’s level of knowledge and skills that affect their ability to interact with the healthcare system (Baker, 2006). A number of valid and reliable measures of HL, such as the Test Of Functional Health Literacy In Adults (TOFHLA) (Parker et al., 1995), have been used to demonstrate that individuals with low HL report worse health status and lower understanding of their medical condition (Ad Hoc Committee on Health Literacy for...
the Council on Scientific Affairs, 1999; DeWalt et al., 2004). The definition of HL includes a clear focus on basic skills, specifically reading and numeracy, rather than on specific knowledge or behaviours related to improving physical health. In contrast, Jorm, et al. (1997) defined MHL as:

“…the ability to recognise specific disorders; knowing how to seek mental health information; knowledge of risk factors and causes, of self-treatments, and of professional help available; and attitudes that promote recognition and appropriate help-seeking.”

(Jorm, et al., 1997, p. 182).

To date, no alternative definition of MHL has been presented in the literature. Within the current definition, Jorm, et al. (1997) identified seven key attributes that constitute MHL, which can be grouped into recognition, knowledge of factors relating to mental health, and attitudes and beliefs about mental disorders. Based on this definition, a clear understanding of an individual’s MHL requires a thorough measurement of all attributes listed.

Jorm, et al. (1997) developed an interview schedule (‘Vignette Interview’) to assess MHL. The Vignette Interview presents a vignette that describes an individual with a mental health difficulty and asks a series of questions relating to participants’ understanding of what is ‘wrong’ with the individual in the vignette (Jorm, et al., 2005; Jorm, et al., 1997). As one of the few measures of MHL, the Vignette Interview has been used extensively. The limitation of this approach is that it is not designed to generate a subscale or total MHL. However, subsequent revisions of the Vignette Interview have allowed for scoring to be conducted and other researchers have developed alternative approaches to address this limitation. These
 developments reflect the need for measures that accurately assess MHL and allow for meaningful comparisons between individuals.

Accordingly, it is important to examine the range of measures currently available that generate a total MHL or subscale score. The benefit of this particular form of measurement is the capacity for increased usability among researchers and practitioners. Similarly, it is important to investigate the extent to which the attributes of MHL, as outlined in the definition of this construct by Jorm, et al. (1997), are captured by current measures. This review sought to examine the frequency with which MHL has been examined in the literature. We undertook a comprehensive review of current measures of MHL that generate a total and/or subscale score to examine their psychometric properties and assess the extent to which these measures test the attributes of MHL as currently defined.

Method

Two searches were undertaken. First, we searched PsychINFO and PubMed for the period 1997 to 2012 using the key term “mental health literacy”. This date was used as MHL was first introduced in 1997. A frequency count was made for total number of publication per year. Second, we identified studies for inclusion in the review of psychometric properties of MHL measures. A comprehensive search of the literature was conducted on PsychINFO and PubMed for the period of 1997 to 2012. Key search terms related to mental health literacy (e.g., mental, health, literacy) and measures (e.g., measure*) were used. The reference lists of the resulting articles were also examined for relevant papers.

Studies were included only if they reported development of a measure of mental health literacy that enabled calculation of either total and/or subscale scores on MHL and contained
extractable data on the development of the measure. Studies that targeted only one category of disorder (e.g., anxiety disorders) were excluded because MHL by definition should encompass knowledge and attitudes relating to a range of mental health disorders and concepts. Figure 1 shows the study selection process.

[APPROXIMATE LOCATION OF FIGURE 1]

Data extraction for psychometrics was based on the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) (Mokkink et al., 2006; Mokkink et al., 2010). The COSMIN is a valid and reliable tool for evaluating the methodological quality of health measurement instruments. It was purpose-designed to provide a clear framework for evaluating psychometric properties of measures and has been used to conduct a number of systematic reviews, including measures of quality of life (Weldam et al., 2012) and employee health (Abma et al., 2011). Studies that met the inclusion criteria were independently scored by two of the authors and scoring discrepancies were discussed and resolved.

Articles were assessed across a range of psychometric properties and scored for quality based on COSMIN guidelines. However, after initial scoring it was found that very few articles reported any measurement properties for the domains assessed, and as such the quality ratings had resulted in a loss of descriptive data. A modified scoring system was therefore introduced, which rated the psychometric properties for each article as being present or absent, rather than providing a rating of quality of reporting. The criteria used for the present or absent ratings were also based on the COSMIN guidelines. In addition to this scoring system, data was extracted regarding the attributes of MHL that were examined in each measure. A study was deemed to
have measured a MHL attribute if the authors explicitly stated measurement of the attribute. This
criteria was implemented to ensure objective assessment between the current raters.

Results

Figure 2 provides an overview of the total number of articles relating to MHL published from
1997 to 2012 and shows the frequency of publications per year. A total of 594 articles were
published within this period. Few publications were present following the proposal of MHL as a
construct in 1997. However, there has been a rapid increase in publications on MHL over the last
eight years, suggesting that this construct is perceived to have utility in the field.

[APPROXIMATE LOCATION OF FIGURE 2]

Thirteen studies were identified as meeting inclusion criteria. Table 1 provides a brief
summary of the key characteristics of the studies, including information regarding participants
and the format used by each measure.

[APPROXIMATE LOCATION OF TABLE 1]

Participants

All measures were tested on adults or young adults, with the exception of Fraser and Pakenham
(2008) who included children. Specific groups including police officers (Compton et al., 2011),
university students (Furnham et al., 2011; Lauber et al., 2005; Smith & Shochet, 2011; Wood &
Wahl, 2006) and emergency personnel (Jordans et al., 2012) were also sampled. Sample sizes
were small to moderate (range = 37 - 587), with the exception of Yap et al. (2012) who used a
large sample (n = 3021). Of the 13 included studies, 6 failed to report sample demographics
including age range (Jorm et al., 2010; Lauber, et al., 2005; Swami et al., 2010) and measures of central tendency (Evans-Lacko et al., 2010; Kitchener & Jorm, 2002; Yap, et al., 2012).

**Format of Measures**

A number of different methodologies were used to assess MHL. The vignette methodology was used by four studies (Jorm, et al., 2010; Kitchener & Jorm, 2002; Smith & Shochet, 2011; Yap, et al., 2012). Unlike the original MHL Vignette Interview (Jorm, et al., 1997), these four studies were constructed to include capacity for total or sub-scale score calculation.

The remaining nine studies used a range of methods including multiple choice questions (Compton, et al., 2011), Likert-response questions (Evans-Lacko, et al., 2010; Wood & Wahl, 2006), dichotomous-response questions (Furnham, et al., 2011; Lauber, et al., 2005; Swami, et al., 2010) and a combination of these methods (Fraser & Pakenham, 2008; Jordans, et al., 2012; Yeap & Low, 2009).

**Measurement Properties**

Results of the structured evaluation of psychometric properties, based on COSMIN guidelines, for each measure are displayed in Table 2.

[APPROXIMATE LOCATION OF TABLE 2]

Studies presented substantial variation in methodological rigour. Across the nine domains outlined in the COSMIN as necessary for high quality methodology, Compton, et al. (2011) reported the highest number of domains assessed (five), followed by Evans-Lacko, et al. (2010) (three). Of the remaining studies, four measured two domains (Jorm, et al., 2010; Smith & Shochet, 2011; Wood & Wahl, 2006; Yeap & Low, 2009), five measured only one domain.
(Fraser & Pakenham, 2008; Furnham, et al., 2011; Kitchener & Jorm, 2002; Swami, et al., 2010; Yap, et al., 2012) and two failed to measure any domains (Jordans, et al., 2012; Lauber, et al., 2005).

The most consistently assessed domains were hypotheses testing (Compton, et al., 2011; Jorm, et al., 2010; Smith & Shochet, 2011; Swami, et al., 2010; Wood & Wahl, 2006; Yap, et al., 2012) and responsiveness (Compton, et al., 2011; Fraser & Pakenham, 2008; Jorm, et al., 2010; Kitchener & Jorm, 2002; Wood & Wahl, 2006). None of the studies assessed measurement error or criterion validity.

**MHL Attributes Measured**

Table 3 displays the MHL attributes measured within each study.

[APPROXIMATE LOCATION OF TABLE 3]

Jorm, et al. (1997) identified seven attributes of MHL in his definition of the construct. None of the studies were scored as having addressed all seven. The most consistently measured attribute was recognition of disorders, which is reported by eight of the studies (Compton, et al., 2011; Evans-Lacko, et al., 2010; Furnham, et al., 2011; Jorm, et al., 2010; Kitchener & Jorm, 2002; Lauber, et al., 2005; Swami, et al., 2010; Yap, et al., 2012). None of the studies assessed either knowledge of how to seek information or knowledge of self-treatment. Compton, et al. (2011) reported the greatest number of attributes measured (four), with the remaining studies measuring either only two attributes (Jorm, et al., 2010) or one attribute (Evans-Lacko, et al., 2010; Furnham, et al., 2011; Kitchener & Jorm, 2002; Lauber, et al., 2005; Smith & Shochet, 2011; Swami, et al., 2010; Wood & Wahl, 2006; Yap, et al., 2012; Yeap & Low, 2009). Two
studies failed to measure any attributes related to MHL (Fraser & Pakenham, 2008; Jordans, et al., 2012), despite clearly identifying their study as an examination of MHL.

In contrast, all studies, with the exception of Swami et al. (2010), reported additional attributes, which were not identified in the definition of MHL by Jorm, et al. (1997). Table 4 displays a summary of the other attributes identified by researchers as central to MHL but not identified by Jorm, et al. (1997).

[APPROXIMATE LOCATION OF TABLE 4]

These attributes included knowledge of a range of factors relating to mental health such as onset and course of a disorder, effectiveness of treatment and study-specific interests such as provision of mental health first aid.

Discussion

The aim of this review was to determine the extent to which MHL has been examined by peer-reviewed studies and to assess the psychometric properties of the currently available measures of MHL that generate a total and/or subscale score. We also sought to assess the extent to which these measures test the attributes of MHL as defined by Jorm, et al. (1997).

A comprehensive search identified over 450 studies that referred to the measurement of MHL, indicating it is a topic of interest of researchers. The frequency count for publications discussing MHL suggests that it is also a topic of steadily growing interest. In view of this interest, it is of concern that the current review findings indicate a striking lack of methodological rigour evidenced in the construction of included measures and little consensus about the attributes of MHL that should be covered when assessing this construct.
Thirteen studies developed a measure of MHL that enabled calculation of total and/or subscale scores and contained extractable data. Of the studies reviewed, each provided a unique methodology for measuring MHL. Studies were grouped into two formats: vignette methodology and mixed methodology including multiple choice, Likert and dichotomous response. Across the majority of studies there was limited psychometric data reported. Furthermore, none of the studies measured all seven attributes listed in the definition of MHL.

Most studies investigating MHL failed to report in sufficient detail information about the sample, questionnaire development and subsequent testing to demonstrate the validity and reliability of their measurement tool. The implication of this finding is that there is no current, quantitative measure of MHL that has been demonstrated to consistently measure MHL across multiple settings and populations.

Measure Format

Examination of available studies indicated there was considerable variation in targeting general population and specific groups, which together with evidence concerning the increasing use of MHL, suggests that MHL is increasingly seen as an important construct that warrants interest across different populations. However, there was variation in the degree to which sample demographics were appropriately reported, with many studies failing to report important information. Similarly, there was considerable variation in the format of the measures suggesting that there is currently limited consensus regarding the most appropriate way to measure MHL. Although the vignette methodology was most commonly used, a number of studies used different formats (e.g., true/false questions and five-point Likert-response), which makes comparison between different studies very difficult.
Since the vignette methodology has been the most widely used means of measuring MHL, it is important to note there are a number limitations to this approach. In its original form, the Vignette Interview did not provide for a total or subscale score, although this has been addressed in other versions of the measure. Further, the wording of the questions in the vignette methodology does not always allow accurate measurement of knowledge, as separate from beliefs, attitudes or opinion. For example, questions such as ‘how likely are particular people to be ‘helpful, ‘harmful’, or ‘neither’ seem more likely to prompt opinions or beliefs about the helpfulness of a particular individual, rather than assessing their knowledge regarding whether that individual has the training or capacity to provide help. The distinction between assessing knowledge or beliefs is important in order to understand the unique contribution of each of these attributes and the potential effect they have on each other. As five of the seven attributes that constitute MHL measure knowledge specifically, it is critical that we are able to separate knowledge from beliefs to make a valid assessment of what an individual knows as opposed to what they believe.

The second limitation with the vignette methodology is that a respondent’s understanding of the mental health difficulty described in the vignette forms the basis for their answers to the remaining questions. For example a participant may be responding to all questions on the assumption that the individual in the vignette is experiencing stress rather than depression. Arguably, these differing assumptions are likely to result in substantially different responses, for example lower endorsement for the use of an antidepressant medication or contact with a psychiatrist. In a national survey, Jorm, et al. (2005) found only 67.3% and 42.5% of participants provided the correct diagnosis of depression or schizophrenia respectively to the vignette presented. Thus, many participants provided answers to the questions that may not have reflected
their knowledge of the aetiology and treatment of these particular disorders. None of the included studies that utilised the vignette methodology in this review had 100% rates of correct recognition and thus may also be similarly affected by this methodological limitation.

**COSMIN Domains Assessed**

None of the studies adequately assessed the nine domains outlined in the COSMIN. The most psychometrically rigorous study (Compton, et al., 2011) addressed only five domains, with the majority addressing only one domain or none at all (Fraser & Pakenham, 2008; Furnham, et al., 2011; Jordans, et al., 2012; Kitchener & Jorm, 2002; Lauber, et al., 2005; Swami, et al., 2010; Yeap & Low, 2009). The lack of thorough assessment has implications for interpretation of results, since there is limited evidence across all studies that what is reported to be measured has been appropriately assessed.

**MHL Attributes Measured**

One of the core components of a successful measure is its ability to accurately assess the various attributes deemed important in the construct. The measure developed by Compton, et al. (2011) is the closest approximation to addressing each of the attributes and is also the most methodologically robust. However, there has been no attempt to systematically assess the attributes of MHL as proposed by Jorm, et al. (1997). Given the lack of measurement of all attributes relating to MHL, it is problematic to conclude that these tools provide a thorough assessment of MHL. Instead, these measures may be measuring only select aspects of MHL. Interestingly, all studies, with the exception of Swami, et al. (2010), include the measurement of additional attributes not included in the definition. Authors have not specified whether the inclusion of additional attributes reflects a belief that the current definition of MHL is
insufficient or whether it simply reflects a weakness in the design of their measure and its ability to assess all MHL attributes.

Strengths and Limitations

One of the strengths of this review was use of the clearly defined set of criteria provided by the COSMIN to assess the psychometric properties of current measures of MHL. We also examined the capacity of these measures to capture the attributes proposed in MHL. As such, this review provides a clear guide to the strengths and limitations of a range of approaches in measuring MHL, which can assist in determining areas for further research. However, there are also several limitations to this current review. Arguably, exclusion of studies that failed to meet our strict inclusion criteria may have resulted in a more limited examination of the current methodologies used to assess MHL. However, only measures of MHL that can generate a total and/or subscale score are likely to provide easy interpretability and increased capacity to compare across individuals.

We applied the COSMIN criteria to all of the studies, even though it could be argued that some of the included studies did not seek to develop and test a new measure and rather, generated a study-specific measure. We also used strict guidelines for determining whether each study measured one of the attributes that define MHL, which may have resulted in an under-reporting of the number of attributes measured. Although practices such as developing a novel measure and using study-specific definitions of attributes may generate interesting findings, ultimately this kind of rigour is inadequate in helping to develop this area of research, which as noted, currently lacks a strong methodological framework.
Given our study strengths, we are confident in recommending that future research should focus on the development of a measure that is psychometrically robust, and ideally, developed in accordance with methodological guidelines such as the COSMIN. Such a measure would include appropriate reporting of validity and reliability. It should also seek to measure all seven of the attributes of MHL. A thoroughly tested measure may be able to provide a more complete understanding of MHL and assist in clarifying the benefit of including, or removing, attributes from MHL to best define and measure this construct. A scorable measure also would allow greater utility in both research and practice, providing the opportunity for rapid and meaningful assessment of an individual’s MHL, for example in use by policy makers to measure trends in MHL and develop appropriate interventions.

Conclusion

Introduction of the construct of MHL has provided an important conceptual development in understanding the factors that impact on an individual’s mental health and help-seeking behaviour. However, comprehensive review of the current measurement of MHL suggests that there is substantial scope for further development and testing of psychometrically robust measures to assess this important construct. Ideally, the development of such a measure would allow us to more fully understand mental health and wellbeing and provide the opportunity to develop early intervention strategies that assist individual’s to maintain their mental health and wellbeing.
References


709 records identified through database searching
4 articles identified through hand searching

458 records after duplicates removed

458 records screened

254 records excluded
- Only one disorder targeted (90)
- Does not generate a total or subscale score (45)
- Not a measure of MHL (3)
- Insufficient information reported (46)
- Utilised a measure already identified (7)

204 full-text articles assessed for eligibility

13 studies used in qualitative synthesis

Figure 1. Flowchart of study selection
Figure 2. Total number of MHL publications between 1997 and 2012
Table 1

Summary of key characteristics of included studies

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Participants</th>
<th>Age Range</th>
<th>N</th>
<th>Format of MHL measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compton, et al. (2011)</td>
<td>Police officers</td>
<td>21-66</td>
<td>N=199</td>
<td>Multiple choice questionnaire</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Study 2: 25-45</td>
<td>M = Not reported</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Study 3: 25-45</td>
<td>SD = Not reported</td>
<td></td>
</tr>
<tr>
<td>Fraser and Pakenham (2008)</td>
<td>Children</td>
<td>12 – 17</td>
<td>N = 44</td>
<td>Likert and dichotomous-response questionnaire</td>
</tr>
<tr>
<td>Furnham, et al. (2011)</td>
<td>University students</td>
<td>18 – 23</td>
<td>N = 426</td>
<td>Dichotomous response questionnaire</td>
</tr>
<tr>
<td>Study</td>
<td>Population</td>
<td>Age Range</td>
<td>Sample Size</td>
<td>Measures</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lauber, et al. (2005)</td>
<td>University students</td>
<td>No age range reported</td>
<td>N = 225</td>
<td>Dichotomous-response questionnaire</td>
</tr>
<tr>
<td>Smith and Shochet (2011)</td>
<td>University students</td>
<td>17-26</td>
<td>N=150</td>
<td>Vignette and Author-generated questionnaire based on Jorm, et al. (1997)</td>
</tr>
<tr>
<td>Swami, et al. (2010)</td>
<td>Adults</td>
<td>No age range reported</td>
<td>N = 477</td>
<td>Dichotomous-response questionnaire</td>
</tr>
<tr>
<td>Wood and Wahl (2006)</td>
<td>University students</td>
<td>16-46</td>
<td>N = 114</td>
<td>Likert-response questionnaire</td>
</tr>
<tr>
<td>Yeap and Low (2009)</td>
<td>Adults</td>
<td>18-60</td>
<td>N = 587</td>
<td>Dichotomous and likert-response questionnaire</td>
</tr>
</tbody>
</table>
Table 2

*Psychometric data, including domains assessed based on the COSMIN*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal consistency</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Measurement Error</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Content validity</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Structural validity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hypotheses testing</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Cross-cultural validity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Criterion validity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Responsiveness</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note:* ‘Yes’ indicates that specific domain was adequately assessed as per COSMIN guidelines.
### Table 3

**MHL Attributes Measured**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of how to seek information</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Knowledge of risk factors</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Knowledge of causes of mental illness</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Knowledge of self-treatment</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Knowledge of professional help available</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Attitudes that promote recognition or appropriate help seeking behavior</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Note.** ‘Yes’ indicates that specific attribute was reported to be assessed
### Table 4

**Additional MHL Attributes Measured**

<table>
<thead>
<tr>
<th>Other MHL attributes measured</th>
<th>Knowledge of the course of the disorder</th>
<th>Help seeking</th>
<th>Beliefs about treatment</th>
<th>Perception of mental health problem in self or family</th>
<th>Knowledge about lifetime prevalence of mental health problem</th>
<th>Knowledge of sex distribution</th>
<th>Psychosocial support</th>
<th>Knowledge of the importance of help seeking for mental illness</th>
<th>Knowledge of the helpfulness of interventions</th>
<th>Knowledge of the helpfulness of availability and affordability of treatment</th>
<th>Help seeking intentions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compton, et al. (2011)</td>
<td>Awareness of mental illness</td>
<td>Confidence in providing mental health first aid</td>
<td>Beliefs about treatment</td>
<td>Social distance</td>
<td>Age of the first episode of each disorder</td>
<td>None</td>
<td>General knowledge</td>
<td>None</td>
<td>General knowledge</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Evans-Lacko, et al. (2010)</td>
<td>Awareness of parental mental illness</td>
<td>Actions taken to implement mental health first aid</td>
<td>Beliefs about treatment</td>
<td>Help provided to others</td>
<td>None</td>
<td>General knowledge</td>
<td>First aid beliefs</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Fraser and Pakenham (2008)</td>
<td></td>
<td>First-aider mental health</td>
<td>Perception of mental health problem in self or family</td>
<td>Beliefs about treatment</td>
<td>Knowledge of lifetime prevalence of mental health problem</td>
<td>Knowledge of the importance of help seeking for mental illness</td>
<td>General knowledge</td>
<td>None</td>
<td>General knowledge</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Furnham, et al. (2011)</td>
<td></td>
<td></td>
<td></td>
<td>Belief in 100% confidence</td>
<td>None</td>
<td>General knowledge</td>
<td>Help seeking</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Jorm, et al. (2010)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>Knowledge of sex distribution</td>
<td>Knowledge of the helpfulness of interventions</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Kitchener and Jorm (2002)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>None</td>
<td>General knowledge</td>
<td>Help seeking</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Lauber, et al. (2005)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>Age of the first episode of each disorder</td>
<td>Knowledge of the helpfulness of availability and affordability of treatment</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Jordans, et al. (2012)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>Knowledge of sex distribution</td>
<td>Knowledge of confidentiality</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Smith and Shochet (2011)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>None</td>
<td>General knowledge</td>
<td>Help seeking</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Swami, et al. (2010)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>Knowledge of lifetime prevalence of mental health problem</td>
<td>Knowledge of the helpfulness of availability and affordability of treatment</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Wood and Wahl (2006)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>Knowledge of sex distribution</td>
<td>Knowledge of confidentiality</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Yap, et al. (2012)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>Age of the first episode of each disorder</td>
<td>Knowledge of the helpfulness of availability and affordability of treatment</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Yeap and Low (2009)</td>
<td></td>
<td></td>
<td></td>
<td>Beliefs about treatment</td>
<td>Knowledge of lifetime prevalence of mental health problem</td>
<td>Knowledge of the helpfulness of availability and affordability of treatment</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>