The Mental Health Literacy Scale (MHLS): A new scale-based measure of mental health literacy.

O'Connor M¹, Casey L².
Abstract

Although Mental Health Literacy (MHL) has been a topic of substantial interest, measurement of this concept has been limited, including a lack of psychometric and methodologically robust scale-based measures of MHL. This study developed a new scale-based measure of MHL, the Mental Health Literacy Scale (MHLS), which assesses all attributes of MHL. Construction of the MHLS was done over three key stages, including measure development, pilot testing and assessment of psychometrics and methodological quality. The resulting measure is a 35 item, univariate scale that is easily administered and scored. Results showed significant differences in scores between mental health professionals and a community sample, as well as individuals with greater experience with mental health, and a significant positive relationship with help-seeking intentions. The MHLS also demonstrated good internal and test-retest reliability. Evaluation of the methodological quality of the MHLS indicated that it has substantial methodological advantages in comparison to existing scale-based measures of MHL. The MHLS can be used in assessing individual and population level differences in MHL and in determining the impact of programs designed to improve MHL.

Keywords: Psychometric, assessment, barriers to treatment, help-seeking
1. Introduction

Mental Health Literacy (MHL) refers to knowledge and attitudes regarding mental health that aid in recognition, management and prevention of mental health issues (Jorm et al., 1997). According to Jorm et al., MHL consists of seven attributes: the ability to recognise specific disorders; knowing how to seek mental health information; knowledge of risk factors and causes; knowledge of self-treatments; knowledge of professional help available; and attitudes that promote recognition and appropriate help-seeking. Although MHL has provided a useful way of conceptualising factors that impact on maintenance of mental health, there are methodological limitations to the measurement of this construct and no existing measure that assesses all attributes of MHL in a scale-based format (O'Connor et al., 2014). The aim of this study was to develop a psychometrically valid, scale-based measure of MHL that assesses all attributes that Jorm et al. (1997) included in his definition of MHL.

There is a low level of knowledge about mental health in the community (Bartlett et al., 2006; Farrer et al., 2008; Jorm et al., 2005), with many individuals unable to identify symptoms of common disorders, such as depression (Jorm et al., 2005) and failing to endorse treatment strategies endorsed by professionals (Jorm et al., 2005; Parker et al., 2001). Consistent with the theory that increased knowledge and positive attitudes are related to better MHL, both nurses and psychiatrists have been found to have higher MHL than lay-people (Caldwell and Jorm, 2000). As higher levels of MHL are related to greater intentions to seek help (Smith & Shochet, 2011), a number of programs have been developed to improve MHL (e.g., Bapat et al., 2009; Kitchener and Jorm, 2002; Potvin-Boucher et al., 2010).
However, there are a number of difficulties with the current measurement of MHL (O'Conner et al., 2014). The most commonly used measure to assess MHL, the Vignette Interview (Jorm et al., 1997), is time-consuming to administer and has no scale-based scoring system. For those scale-based measures that have been developed, only limited psychometric data has been reported and none of these measures assess all of the attributes of MHL (O'Connor et al., 2014).

The aim of this study was to use a comprehensive methodological approach to develop a new scale-based measure of MHL, the Mental Health Literacy Scale (MHLS), to more easily and accurately assess an individual’s level of MHL. This assessment has implications for determining in which areas individuals may require further support, and in evaluating the effectiveness of interventions intended to improve MHL. The goal was to develop a psychometrically and methodologically strong measure that would enable assessment of all attributes of MHL.

2. Method

Figure 1 outlines the three phases of development: measure development, pilot testing and assessment of psychometrics and methodological quality.
Figure 1 – Flowchart for the development of the MHLS
2.1 Measure development – phase one

Operational definitions of the seven attributes of MHL were developed using an iterative process until a consensus was reached within the clinical panel, which included clinical psychology staff and the research team. Feedback suggested that there is an insufficient level of knowledge in the field to enable differentiation of risk factors for mental illness and causes of mental illness, resulting in these attributes being merged into one attribute; knowledge of risk factors and causes. Table 1 presents these operational definitions.

Table 1

Operational definitions of MHL attributes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Operational definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to recognise specific disorders</td>
<td>Ability to correctly identify features of a disorder, a specific disorder or category of disorders</td>
</tr>
<tr>
<td>Knowledge of how to seek mental health information</td>
<td>Knowledge of where to access information and capacity to do so</td>
</tr>
<tr>
<td>Knowledge of risk factors and causes</td>
<td>Knowledge of environmental, social, familial or biological factors that increase the risk of developing a mental illness</td>
</tr>
<tr>
<td>Knowledge of self-treatments</td>
<td>Knowledge of typical treatments recommended by mental health professionals and activities that an individual can conduct</td>
</tr>
<tr>
<td>Knowledge of professional help available</td>
<td>Knowledge of mental health professionals and the services they provide</td>
</tr>
<tr>
<td>Attitudes that promote recognition and appropriate help-seeking</td>
<td>Attitudes that impact on recognition of disorders and willingness to engage in help-seeking behaviour</td>
</tr>
</tbody>
</table>
2.1.1 Item development.

Items were generated for each of the six attributes by the research team and clinical panel. Items requiring a correct answer were checked by consulting the relevant literature for consensus, and through additional discussion with the clinical panel. Table 2 outlines key decisions regarding item development.
### Table 2

#### Rationale for item development

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Development rationale</th>
<th>Response format</th>
<th>Number of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition of disorders</td>
<td>Items had a stronger focus on the most common disorders (based on data from the Australian Bureau of Statistics (2007)). Descriptions of disorders were based on Diagnostic and Statistical Manual of Mental Disorders IV TR (American Psychiatric Association, 2000) criteria and grouped into vignette items and specific diagnostic items. No items were affected when the Diagnostic and Statistical Manual of Mental Disorders transitioned to 5th edition</td>
<td>Multiple choice question</td>
<td>20</td>
</tr>
<tr>
<td>Knowledge of how to seek mental health information</td>
<td>Items were adapted from the Vignette Interview (with permission, A. Jorm, personal communication, September 27, 2012). Items were also included to assess an individual’s capacity to access mental health information, comparable to the approach for measuring Health Literacy (Baker, 2006). The format of capacity items was modelled on the Patient Activation Measure (Hibbard et al., 2004)</td>
<td>Multiple choice question and Likert (4-point scale)</td>
<td>12</td>
</tr>
<tr>
<td>Knowledge of risk factors and causes</td>
<td>Items assessed knowledge of risk factors for developing mental illness, including a number of common misconceptions about risk factors. Items were also developed assessing knowledge of common at-risk groups, which were based on Australian Bureau of Statistics (2007) data</td>
<td>Dichotomous: True/False</td>
<td>14</td>
</tr>
<tr>
<td>Knowledge of self-treatments</td>
<td>Items were developed based on the clinical experience of the clinical panel and included knowledge of common strategies typically recommended by mental health practitioners to improve mental health and wellbeing</td>
<td>Multiple choice question</td>
<td>7</td>
</tr>
<tr>
<td>Knowledge of professional help available</td>
<td>Items were developed based on the clinical experience of the clinical panel and included knowledge of the services typically provided by mental health practitioners</td>
<td>Multiple choice question</td>
<td>9</td>
</tr>
<tr>
<td>Attitudes that promote recognition and appropriate help-seeking</td>
<td>Items were adapted from the Vignette Interview (with permission, A. Jorm, personal communication, September 27, 2012) and similarly-worded additional items were included based on feedback from the panel</td>
<td>Likert (5-point scale)</td>
<td>17</td>
</tr>
</tbody>
</table>
2.2 Item testing – phase two

The MHLS-Pilot (MHLS-P) consisted of 79 questions, which were provided to an additional panel of practicing mental health professionals ($n=7$) as well as to the original clinical panel for feedback. The MHLS-P was then administered to a community sample ($n=202$) in order to conduct a preliminary analysis of items.

2.2.1 Participants and procedure

Participants were men ($n=62$) and women ($n=140$) recruited online through a snowballing process using social media (facebook, twitter, email). Participants were required to be 18 years or older and Australian residents. The mean age of the sample was 33.25 years ($SD=16.02$), with the majority of participants being Caucasian (91.8%), possessing at least a Bachelor’s degree (54%) and located in a major city (79.7%). Participants were provided with a link to the study to complete the MHLS-P online.

2.2.2 Results

Data analysis was conducted using the Statistical Package for the Social Sciences (SPSS) version 17.0. For dichotomous items, 39 items demonstrated ceiling effects, having higher than 80% correct endorsement rate. According to Rummel (1970), dichotomous variables with a 90-10 split should be considered for removal. To ensure the MHLS would be a sensitive measure of MHL, a more conservative split of 80-20 was used. In total, 39 items were in excess of the cut-off. Likert items demonstrated considerable skewing and kurtosis, indicating that participants had strong positive beliefs in their capacity to seek mental health information and attitudes toward mental health difficulties.
Based on these results, a number of changes to items were implemented and are presented in Table 3.

Table 3

Summary of results and modifications

<table>
<thead>
<tr>
<th>Statistical analyses</th>
<th>Modifications</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ceiling effects on items</td>
<td>Excluded 28 items. Retained those items with less than 80% correct response rate or where theoretically relevant to include assessment of all MHL attributes</td>
<td>Improves the sensitivity of the measure and ability to differentiate participants with high MHL and low MHL</td>
</tr>
</tbody>
</table>

Feedback from clinical panel

Include reverse scored items

False items included that can be reverse scored. Many false items were developed by using the most frequently endorsed incorrect answer

Need consistency across response format

All knowledge items (MCQ and dichotomous) adapted to a 4-point Likert format and the structure of questions changed to include ‘to what extent do you think’

Increases the difficulty of the measure as answers are less obvious and it requires greater cognitive processing (Idaszak and Drasgow, 1987)

Allows a greater amount of information to be obtained for each question. Items were written such that there was still a correct answer and that a participant’s response on a 4-point Likert scale indicated their degree of knowledge

2.3 Assessment of psychometrics and methodological quality – phase three

Following the statistical analyses and feedback from Phase Two, the MHLS-P was refined (MHLS-Pilot-Revised, MHLS-P-R), resulting in a total of 51 items, which consisted of ability to recognise disorders (21), knowledge of where to seek information (4), knowledge of risk factors and causes (2), knowledge of self-treatment (2), knowledge of professional help available (5) and attitudes that promote recognition or appropriate help-seeking behaviour (17).
The number of items across attributes differed, as some attributes required a larger number of items to appropriately address that attribute (e.g., recognition of disorders). In order to further refine the measure and assess its psychometrics, the MHLS-P-R was administered to a community sample \((n=372)\) and sample of mental health professionals \((n=43)\).

### 2.3.1 Participants

#### 2.3.1.1 Community sample.
Ninety-four male and 278 female first year university students completed this study as part of their course credit obtained in psychology courses. Testing was conducted early in the first semester in order to minimise the impact that undertaking courses in psychology may have had. All participants could enter the draw for a small prize, drawn at the conclusion of the research.

The mean age of participants was 21.10 years \((SD=6.27)\). Most participants were Caucasian (73.7%), had the highest level of qualification as a Secondary School certificate (76.1%) and lived in a major city (73.7%).

#### 2.3.1.2 Mental health professionals.
Thirty-seven female and six male mental health professionals completed the study. The mean age of participants was 33.09 years \((SD=8.01)\). The majority of participants were Caucasian (86.0%), had completed a Bachelor’s degree or higher (95.4%) and lived in a major city (88.4%). Participants were recruited through professional networks of the researchers.

### 2.3.2 Materials

#### 2.3.2.1 Demographics.
Demographic information collected included age, gender, ethnicity, education and residence.
2.3.2.2 Mental Health Literacy Scale Pilot-Revised (MHLS-P-R). The measure consisted of 51 items, with 25 reverse scored items. Items were scored according to the value selected (i.e., on a 1-4 scale, if a participant selected two, they scored two points). Scores for the scale were determined by summing items.

2.3.2.3 Mental health experience. Items assessing the participant’s level of experience with mental health, including experience of a mental illness and access to treatment were collected from the community sample only.

2.3.2.4 The General Help-seeking Questionnaire (GHSQ). The GHSQ is a measure of intention to seek help from different sources (Wilson et al., 2007). The question used in order to assess general help-seeking was, “Imagine you are experiencing a personal-emotional problem or mental health difficulty”. Response options are scored on a Likert scale ranging from ‘extremely unlikely’ (1) to ‘extremely likely’ (7) and include a range of possible sources for accessing help. A higher score indicates greater intentions to seek help. Items can be combined to form a total, formal sources of help and informal sources of help scales. The GHSQ is significantly correlated with actual help-seeking behaviour, for example use of informal sources, such as help from an intimate partner ($r_s=.48, p<.001$) and access to counselling ($r_s=.17, p<.05$) (Wilson et al., 2007). The GHSQ also demonstrates good test-retest reliability ($\alpha=.92$) (Wilson et al., 2007).

2.3.2.5 Kessler Psychological Distress Scale 10 (K10). The K10 is a measure of general psychological distress (Kessler et al., 2002). It is a 10 item scale, with participants indicating their level of agreement to items on a 5-point Likert scale ranging from none of the time (1) to all of the time (5), with a higher score indicating greater levels of distress. An example question is, “During the last 30 days, about how often did you feel tired out for no good reason?”. The K10
has demonstrated strong discriminant validity in its ability to distinguish between clinical and non-clinical populations (Andrews and Slade, 2001; Australian Bureau of Statistics, 2007) and displays good internal consistency ($\alpha = .93$) (Kessler et al., 2002).

2.3.3 Procedure

Prior to the recruitment of participants, ethical approval was obtained from Griffith University Ethics Committee. A link to an online version of the MHLS-P-R using Limesurvey was provided. Presentation of items within the six attributes was randomised to reduce order effects (Jon and Alwin, 1987).

In order to assess test-retest reliability, participants in the community sample were invited to complete the measure again two weeks later. To link responses from the first and second administration, participants generated a unique code known only to them, which would allow their response to be compared. The administration of the retest included the same instructions and items for the MHLS-P-R.

3. Results

3.1 Assessing factorability

An initial principal axis factoring was conducted on 51 items using an oblique rotation (direct oblimin). Use of parallel analysis suggested extraction of 9 factors, while the scree plot and examination of factors that contributed at least 5% of the variance (Hair et al., 2009) suggested a structure with 3 factors. Analysis of the solution was conducted with extraction of 9 and 3 factors. In both solutions, communalities were low, suggesting that the proportion of variance in items explained by the derived factors was low, with mean factor loadings of .166
and .239 respectively. Thus the results indicated that a univariate structure was the most statistically and theoretically appropriate.

3.2 Item reduction and reliability

In order to reduce the number of items in the measure and improve reliability of the MHLS, items with a corrected item-total correlation less than 2 were deleted sequentially to improve the overall Cronbach’s alpha. This value was selected by examining the item set and determining the most appropriate cut-off, which resulted in a reduction of items without removing too many items. The alpha level following the removal of 22 items was .879, which also represented the highest possible Cronbach’s alpha, as indicated by examining the projected Cronbach’s alpha if additional items were deleted. In order to retain at least one item to assess each of the attributes present in MHL, 6 items were re-entered, resulting in a total of 35 items and a final alpha level of .873.

The factor structure of the final 35 item measure was re-analysed following the same procedure as the original factor analysis. The most viable structure was a 4 factor structure, however, similar to the prior analysis, there were low communalities and mean factor loadings (.251). This indicated that the univariate structure still provided the most meaningful interpretation.

To establish the reliability of the measure, participants were retested two weeks after their initial completion of the MHLS, with the results showing good reliability ($r(69) = .797$, $p<.001$). Standard error of measurement was also calculated and found to be 5.70.

The final version of the MHLS included a total of 35 items, which consisted of ability to recognise disorders (8), knowledge of where to seek information (4), knowledge of risk factors
and causes (2), knowledge of self-treatment (2), knowledge of professional help available (3) and attitudes that promote recognition or appropriate help-seeking behaviour (16). All psychometric testing was conducted using these items in the MHLS.

3.3 Descriptives for the MHLS

The community sample was used to generate descriptives for the MHLS. Mean score for the scale was 127.38 (SD=12.63, Minimum=92.00, Maximum=155.00, 95% CI=126.09 to 128.67). Overall, the scale was somewhat normally distributed (Skewness= -.115, Kurtosis= -.231). There were no missing responses in the data.

3.4 Known groups assessment

A series of independent sample t-tests were conducted to examine the differences between groups expected to differ in their MHL. Due to the number of t-tests conducted, a Bonferroni correction was applied, resulting in a significant alpha level of 0.01. Mental health professionals had significantly higher MHL (M=145.49, SD=7.19) than the community sample (M=127.38, SD=12.63). Levene’s test for equality of variances indicated unequal variances (F=13.195, p<.001) so degrees of freedom were adjusted, t(76.21) = -14.18, p<.001. The magnitude of the difference in the means (mean difference= -18.11, 95% CI: -20.65 to -15.57) was large (d=1.76).

Further analyses showed that individuals who have had a mental illness had significantly higher MHL (M=130.97, SD=13.21) than those who had not (M=125.19, SD=11.76), t(370) = 4.39, p<.001. The magnitude of the difference in the means (mean difference= 5.79, 95% CI=3.19 to 8.38) was small (d=0.46). Individuals who had been to see a mental health practitioner had significantly higher MHL (M=133.53, SD=12.02) than those who had not
The magnitude of the difference in the means (mean difference = 9.65, 95% CI = 7.15 to 12.14) was large (d = 0.82). Finally, individuals who had a family member or friend with a mental illness had significantly higher MHL (M = 129.53, SD = 12.12) than those that did not (M = 122.69, SD = 12.49), t(370) = 5.00, p < .001. The magnitude of the difference in the means (mean difference = 6.84, 95% CI = 4.15 to 9.52) was medium (d = 0.56).

3.5 Construct validity

The MHLS was significantly positively correlated with the GHSQ total scale r(370) = .234, p < .001, the GHSQ formal scale r(370) = .146, p = .005 and informal scale, r(370) = .185, p < .001, indicating that individuals with higher MHL are more likely to seek help overall, and from formal and informal sources. There was no significant relationship between MHL and the K10 r(370) = -.087, p < .092 ns, indicating that levels of psychological distress are not related to levels of MHL.

3.6 Consensus Based Standards for the Selection of Health Instruments (COSMIN)

The methodological quality of the MHLS was examined using the COSMIN (Mokkink et al., 2006; Mokkink et al., 2010). Domains were determined as being adequately assessed if they included the minimum level of information required under COSMIN criteria, a methodology which was utilised by O'Connor et al. (2014).

In total, six of the nine domains were determined as being adequately assessed: internal consistency, reliability, measurement error, content validity, structural validity, and hypotheses testing. Criterion validity could not be assessed due to the absence of an accepted gold-standard
for scale-based measurement of MHL. Cross-cultural validity and responsiveness are being assessed in a separate study currently underway.

4. Discussion

Although the definition of MHL developed by Jorm et al. (1997) has been widely used, there has been no systematic attempt to develop a psychometrically robust instrument using this definition (O’Connor et al., 2014). The aim of this study was to use a comprehensive methodological process to develop a scale-based measure of MHL that assessed all attributes of MHL.

This process resulted in a 35 item questionnaire that can be used to assess knowledge and attitudes of a range of areas in mental health. Use of this scale will enable efficient identification of individuals who have low levels of MHL and may benefit from further education or support. The MHLS will also allow the detection of changes within an individual in order to assess the impact of programs to improve MHL.

The development of the MHLS was an iterative process, which included the extensive use of feedback from the clinical panel, construction of operational definitions to guide item development for each attribute and several phases of item testing and review. The rigour of this process is reflected in the psychometric properties that the MHLS displays. The univariate nature of the MHLS, indicating that MHLS scores capture a combination of all attributes of MHL, provides a theoretically meaningful structure that is consistent with the definition of MHL.

The MHLS demonstrated good internal and test-retest reliability, and good validity. In line with previous research, mental health professionals had significantly greater MHL (Caldwell and Jorm, 2000), as did individuals who had greater direct or indirect experience with mental
illness (Furnham et al., 2011; Lauber et al., 2005). Scores on the MHLS were also significantly correlated with help-seeking intentions (Smith and Shochet, 2011). The MHLS was not correlated with psychological distress, suggesting that the relationship between help-seeking and MHL was not influenced by levels of distress.

4.1. Limitations

There are a number of limitations that are important to note. The community sample used in the final phase of psychometric testing consisted of first year university students undertaking psychology courses. Arguably, these students may have a range of characteristics that could inflate their MHL relative to a more representative community sample. However, there was considerable variability in scores across the MHLS, indicating that the MHLS is sensitive to detecting differences within a sample, and further, their scores showed significant differences compared to the mental health professionals.

The goal of developing a brief and easily administered measure of MHL may have resulted in insufficient assessment of the identified attributes of MHL. However, research into mental health has been characterised by difficulties in achieving consensus on many questions (e.g., Luborsky et al., 2002). Consequently, there are many possible ways that each attribute may have been interpreted and assessed. To address this issue, multiple sources were used to achieve consensus in guiding item development and testing, including use of the literature, the clinical panel and a comprehensive testing process including two stages of testing to refine the items.

4.2 Implications and future research

The MHLS provides the first scale-based measure to assess all attributes of MHL. It has good psychometric properties and is easily administered and scored. Consequently, the MHLS
offers considerable benefits in research and practice. Use of the MHLS will aid efficient
evaluation of programs that aim to improve MHL and ensure they are adequately addressing all
attributes of MHL.

In addition to our ongoing research into the cross-cultural validity of the MHLS, future
research could usefully extend the generalizability of the current findings by assessing the
psychometric properties of the MHLS with other samples, with the aim of developing
statistically robust norms that can be used to guide future use of the MHLS. This level of
research is important as it has the capacity to identify particular groups who may require further
support in developing their MHL and to support policy development in this important area.

Acknowledgement

The authors thank Dr Janine Lurie for her assistance with statistical analyses.

Declaration of interests

No conflicts of interests


