

Psychometric Evaluation of the Malay Version of the Mental Health Knowledge Schedule (MAKS-M): Evidence of Structural and Cultural Limitations



Vie Cheong Thong¹, Nicholas Tze Ping Pang¹, Assis Kamu^{2,*}, Cheryl Yee Ying Phang¹, Alvin Daniel Brata Laurio¹, Padmashini Mariappan¹, Sukhashinee Richard Silvaa¹, Nur Khaleeda Nadzwa Kamri¹, Mohd Faizal Madrim¹, Azizan Omar¹ and Chong Mun Ho²

¹Faculty of Medicine and Health Sciences, University Malaysia Sabah, Kota Kinabalu, Malaysia

²Faculty of Science and Natural Resources, University Malaysia Sabah, Kota Kinabalu, Malaysia

Abstract:

Introduction: Mental health literacy instruments are often adapted across cultures with the assumption that their constructs remain stable. This study critically evaluates the psychometric performance of the Malay version of the Mental Health Knowledge Schedule (MAKS-M) to examine whether its original structure transfers to the Malaysian context and to identify factors contributing to cross-cultural underperformance.

Methods: A cross-sectional survey was conducted among 203 university-based respondents. Reliability indices, confirmatory factor analysis (CFA), and convergent validity testing were used to assess structural performance.

Results: Both subscales showed poor internal consistency (knowledge: $\alpha = 0.262$; attitude: $\alpha = 0.236$) and weak item-rest correlations, with reversed items performing particularly poorly. CFA indicated suboptimal model fit (CFI = 0.712, RMSEA = 0.124), with several reversed items showing non-significant or negative factor loadings. Convergent validity was partially supported through significant correlations with the BIPM interest subscale.

Discussion: The MAKS-M demonstrated significant psychometric limitations, likely due to linguistic ambiguities, cultural incongruities, and conceptual overlap between knowledge and attitude constructs. In particular, reversed items appear especially vulnerable to linguistic and cultural misinterpretation, and the results suggest that the MAKS factor structure may not be inherently stable when applied outside Western sociocultural contexts.

Conclusion: Rather than confirming validity, the study identifies structural and cultural barriers to the reliable use of the MAKS in Malaysia. The findings highlight broader methodological challenges in direct cross-cultural scale adaptation.

Keywords: Mental health literacy, Attitude toward mental illness, Scale validation, Psychometrics, Cultural adaptation, Malaysia.

© 2026 The Author(s). Published by Bentham Open.

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

*Address correspondence to this author at the Faculty of Science and Natural Resources, University Malaysia Sabah, Kota Kinabalu, Malaysia; E-mail: assis@ums.edu.my

Cite as: Thong V, Pang N, Kamu A, Phang C, Laurio A, Mariappan P, Silvaa S, Kamri N, Madrim M, Omar A, Ho C. Psychometric Evaluation of the Malay Version of the Mental Health Knowledge Schedule (MAKS-M): Evidence of Structural and Cultural Limitations. Open Psychol J, 2026; 19: e18743501444062. <http://dx.doi.org/10.2174/0118743501444062260623055654>



Received: October 16, 2025

Revised: February 23, 2026

Accepted: March 11, 2026

Published: July 02, 2026



Send Orders for Reprints to
reprints@benthamscience.net

1. INTRODUCTION

Mental health literacy refers to the knowledge and beliefs concerning mental disorders that facilitate their identification, management, or prevention [1]. It is a crucial factor in reducing stigma and encouraging timely help-seeking behaviors [2]. Evaluating the public's understanding of and attitudes toward mental health is therefore fundamental to the development of effective educational initiatives and anti-stigma interventions [3]. The Mental Health Knowledge Schedule (MAKS), created by Evans-Lacko and colleagues (2010), is a concise self-report tool designed to measure two main aspects of mental health literacy: objective knowledge and stigma-related attitudes [4].

Although the MAKS has been used in various countries, its application outside English-speaking contexts requires careful translation and cultural adaptation [5, 6]. Cross-cultural validation is vital because cultural norms [7], language nuances [8], and conceptualizations of mental health [9] can influence how respondents interpret scale items, potentially affecting measurement reliability and validity [10]. In Malaysia, mental health literacy research is growing, but there remains a lack of rigorously validated Malay-language tools for measuring both knowledge of and attitudes toward mental health [11, 12]. Mental health literacy is not a purely cognitive construct in the Malaysian context; interpretations are shaped by cultural norms, moral beliefs, and religious meaning systems. As such, instruments developed in Western settings require conceptual adaptation, not solely translation.

The MAKS consists of 12 items: six assessing knowledge of mental health-related facts and conditions and six assessing stigma-related attitudes toward mental health [4]. Its brevity and ease of administration make it attractive for large-scale surveys, but concerns have been raised regarding the performance of certain items—particularly reversed items—which have been shown to introduce response biases and reduce internal consistency [12].

Previous validation studies have yielded mixed results. For example, studies in European settings have generally supported the two-factor structure but reported modest internal consistency, especially for the knowledge subscale [4]. Research in Asian contexts, such as China and Singapore, has highlighted additional challenges: linguistic translation may alter item meaning, reversed items may cause confusion, and ceiling effects may occur in populations with high baseline familiarity with mental health topics [13, 14].

In Malaysia, few instruments have been validated for the simultaneous assessment of mental health knowledge and stigma-related attitudes. Existing studies have often relied on unvalidated or partially adapted tools, limiting the comparability and interpretability of findings [15, 16]. A validated Malay version of the MAKS could fill this gap, enabling consistent measurement across research, policy

evaluation, and community education initiatives. However, any adaptation must be rigorously tested to ensure that it maintains conceptual equivalence, cultural relevance, and robust psychometric properties [6, 17, 18].

Despite the widespread use of mental health literacy instruments across cultural contexts, most validation studies implicitly assume that the underlying constructs and factor structures of the original scales are stable and transferable [17]. Consequently, cross-cultural adaptations often focus on improving model fit or internal consistency rather than explicitly testing whether the original construct architecture remains valid in non-Western settings. When instruments such as the MAKS are adopted uncritically, there is a risk that observed scores may reflect measurement artefacts rather than true differences in mental health knowledge or stigma-related attitudes, potentially leading to misleading research conclusions and policy decisions [17, 18].

Accordingly, the present study is not positioned as a conventional validation exercise aimed at endorsing the MAKS-M for applied use. Instead, it is framed as a critical psychometric evaluation that examines whether the theoretical structure, item design, and measurement assumptions of the original MAKS remain valid when transferred to a Malaysian cultural and linguistic context. By documenting points of psychometric failure rather than obscuring them, this study contributes methodological insight into the limits of direct scale adaptation in cross-cultural mental health research.

2. METHODOLOGY

This research focuses on the translation, cultural adaptation, and validation of the Mental Health Knowledge Schedule (MAKS-M) for application among Malay-speaking individuals in Malaysia. Standardized international guidelines for cross-cultural adaptation were followed, after which several psychometric evaluations were conducted to assess the scale's reliability and construct validity [18]. An overview of the study process is presented in the flowchart (Fig. 1).

2.1. Translation Process

The translation process involved both the original English version and the previously tested MAKS-M by Pheh *et al.* (2017) [19].

The original MAKS, developed in English, was translated into Malay following a standardized forward-backward translation procedure based on WHO guidelines. A bilingual expert independently forward-translated the scale into Malay. A separate bilingual translator, blinded to the original version, then back-translated the Malay version into English.

The Mental Health Knowledge Schedule-Malay Version (MAKS-M) is an adapted version of the original 12-item MAKS that was previously translated and underwent preliminary validation by Pheh *et al.* (2017) [19]. The MAKS-M was further evaluated by having another team of translators back-translate it into English.

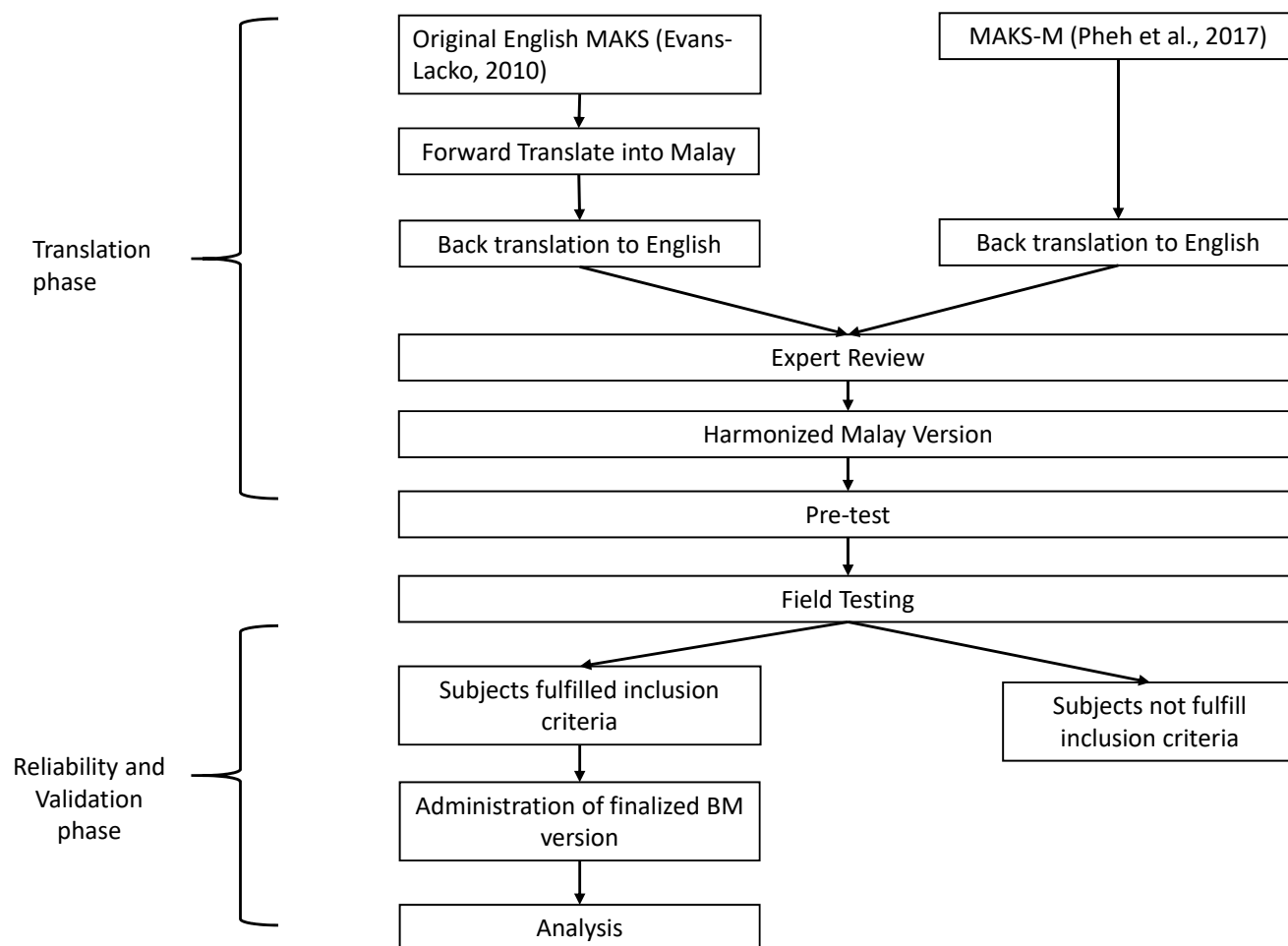


Fig. (1). Translation and validation flowchart.

Both versions were compared and reviewed by a panel of three experts in both the linguistic and clinical fields. During the review, some minor changes were made to the language for better clarity, such as the addition of descriptions to Item 10, “bipolar disorder” (English version), and Item 12, “grief” (English version), by providing direct Malay translations with descriptions for better clarity. The word “problems” in the translation was also replaced with “symptoms” in Malay to reduce stigma in language delivery. The overall context and structure of the scale remained unchanged.

The final translated version is shown in Appendices **A** and **B**.

2.2. Pre-test and Field Testing Process

A preliminary test with 20 bilingual university students was carried out to evaluate the face validity and clarity of the translated MAKS items. Feedback from this group led to minor linguistic revisions that enhanced the readability of the Malay version. These individuals were excluded from the main validation sample, and their responses were used solely for refinement before broader field testing.

For field testing, a cross-sectional study design was adopted to examine the psychometric characteristics of the Malay Mental Health Knowledge Schedule (MAKS-M). Data were obtained from 203 participants enrolled at a public university in Malaysia and recruited through convenience sampling between 1 June 2025 and 10 August 2025. Participants received study information and provided informed consent before completing the questionnaires. No personal identifiers were collected, and once submitted, responses could not be altered or resubmitted.

Respondents first completed the Malay version of the MAKS, followed by the Balanced Index of Psychological Mindedness (BIPM), and finally the original English version of the MAKS. The recruitment of bilingual students ensured language proficiency and familiarity with self-administered questionnaires, allowing for reliable comparisons between the Malay and English versions. This procedure is consistent with common practices in preliminary validation research designed to assess psychometric performance [6].

While this design allows for a preliminary assessment of conceptual equivalence, it may also introduce priming or contamination effects, whereby exposure to one language version influences responses to the other. Accordingly, bilingual administration may be acknowledged as a methodological limitation rather than a strength, and the findings should be interpreted with caution.

2.3. Inclusion and Exclusion Criteria

Inclusion criteria for participation were as follows: (i) aged 18 years and above; (ii) literate in both Malay and English; (iii) currently enrolled as university students; (iv) no reported history of major psychiatric or neurological disorders (self-declared); and (v) provided informed consent to participate in the study.

Participants were excluded if they: (i) had insufficient proficiency in either Malay or English to complete the study questionnaires reliably; (ii) reported a current or past diagnosis of a severe psychiatric disorder (*e.g.*, psychosis, bipolar disorder) that could impair comprehension or participation; or (iii) had previously participated in the pilot/face validity testing phase of this study.

All participants were non-clinical volunteers.

2.4. Sample Size

Sample size determination was guided by established psychometric guidelines for validation studies. A minimum subject-to-item ratio of 10:1 was set to ensure adequate statistical power for internal consistency analysis and confirmatory factor analysis (CFA), as recommended in scale development literature [20, 21]. Additionally, a medium effect size (Cohen's $d = 0.5$), which is widely accepted as the minimum threshold for significance in psychometric research, was used as a benchmark for sample adequacy.

The Malay version of the Mental Health Knowledge Schedule (MAKS-M) comprises 12 items divided into two subscales: (i) Subscale A (Mental Health Knowledge): 6 items; (ii) Subscale B (Stigma-related Attitudes): 6 items.

Based on the minimum ratio of 10 participants per item, the required sample size was calculated as:

Required minimum sample size = Number of items $\times$ 10 = $12 \times 10 = 120$ participants

A total of 203 participants were successfully recruited, which exceeds the recommended minimum and provides a stable basis for evaluating internal consistency (*e.g.*, Cronbach's α , McDonald's ω), item-total correlations, and the factorial structure through CFA. This sample size also ensures sufficient power for testing convergent validity with related constructs such as psychological mindedness [20].

2.5. Measurements

2.5.1. Mental Health Knowledge Schedule (MAKS)

The Mental Health Knowledge Schedule (MAKS) is a validated instrument developed by Evans-Lacko *et al.* (2010) to assess stigma-related knowledge about mental

health among the general public. It was designed to address the lack of psychometrically sound tools for evaluating mental health literacy, particularly within the context of stigma-reduction interventions [4].

The MAKS consists of 12 items divided into two main sections: Items 1 to 6 assess general stigma-related knowledge across six domains: help-seeking, recognition, support, employment, treatment, and recovery; Items 7 to 12 evaluate participants' recognition of specific conditions (*e.g.*, depression, schizophrenia) as mental health problems, providing context for interpreting conceptualizations of mental illness.

Respondents rate their agreement with each statement on a 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5), with an option for "don't know" coded as neutral (3). Several items are reverse-scored to reflect correct knowledge direction.

Psychometric evaluation demonstrated that the MAKS has moderate internal consistency (Cronbach's $\alpha = 0.65$ for items 1-6) and good test-retest reliability (Lin's concordance coefficient = 0.71), with individual item kappa values ranging from 0.57 to 0.87. While the tool was not designed to function as a unidimensional scale, it captures diverse knowledge components relevant to stigma, making it suitable for item-level analysis [4].

2.5.2. Balanced Index of Psychological Mindedness - Malay Version (BIPM-M)

The BIPM-M is a 14-item self-report instrument adapted from the original BIPM to assess individuals' insight and interest in psychological processes. The scale was translated into Bahasa Malaysia following WHO guidelines, involving forward- and backward-translation by language and content experts, followed by pre-testing with 50 native speakers. A cross-sectional validation study involving 141 university students was conducted to evaluate its psychometric properties [22].

Factor analysis confirmed a two-factor structure (Insight and Interest), consistent with the original scale. The BIPM-M demonstrated excellent internal consistency (Cronbach's $\alpha = 0.87$) and strong reliability (McDonald's $\omega = 0.88$; GLB = 0.95). It showed moderate correlations with AAQ-II and MAAS, supporting its concurrent validity. The BIPM-M is a culturally appropriate and psychometrically sound tool for measuring psychological mindedness among Malay-speaking populations [22].

2.6. Data Analysis Methods

The data analysis to validate the Malay version of the Mental Health Knowledge Schedule (MAKS-M) was conducted systematically to establish the instrument's reliability and validity. To begin, internal consistency was assessed using Cronbach's alpha, which evaluates the degree to which items within a scale are interrelated [23]. Each of the two subscales-Mental Health Knowledge (items 1 to 6) and Stigma-related Attitudes (items 7 to 12)-was analyzed independently to ensure that the translated items consistently measured their respective constructs. Higher Cronbach's alpha values indicate greater internal

consistency, with values above 0.70 generally considered acceptable [23].

In addition to Cronbach's alpha, McDonald's omega and the Greatest Lower Bound (GLB) reliability coefficients were computed for both subscales. McDonald's omega accounts for varying factor loadings across items, offering a more precise estimate of reliability in cases where the assumption of equal item contribution (tau-equivalence) may not be met [24]. GLB, on the other hand, provides a lower-bound estimate of true reliability and is especially useful in psychometric evaluations when item true-score variances differ. A GLB value exceeding 0.80 is typically indicative of strong internal consistency and adds further confidence in the scale's reliability. Corrected item-total correlations, often referred to as item-rest correlations, were computed to evaluate reliability at the item level. This statistic indicates the relationship between a single item and the total score of the remaining items within the subscale, thereby showing the extent to which each item corresponds to the construct being measured [24].

Construct validity was assessed by examining both convergent and discriminant validity. Convergent validity was evaluated by computing Spearman's correlation coefficients between the MAKS-M subscales and the subscales of the Balanced Index of Psychological Mindedness (BIPM), namely Interest and Insight. These constructs are theoretically related to mental health knowledge and openness to psychological understanding, and moderate to high positive correlations were expected. Discriminant validity was determined by comparing correlations between the MAKS-M and constructs assumed to be theoretically unrelated. This approach allowed for an evaluation of the scale's ability to differentiate between related and unrelated constructs. The use of Spearman's rho was justified due to the non-normal distribution of the subscale scores, as indicated by both the Kolmogorov-Smirnov and Shapiro-Wilk tests.

To further examine cross-language construct equivalence, a subset of bilingual participants completed both the English and Malay versions of the MAKS. Spearman correlations were calculated between subscale scores across the two versions. High correlation coefficients would suggest that the translated version retains conceptual and measurement equivalence with the original English scale, thereby supporting its use in the local context [25, 26].

Confirmatory factor analysis (CFA) was conducted to test the transportability of the original two-factor structure proposed by Evans-Lacko *et al.*, rather than to optimise model fit within the Malaysian sample. As the dimensionality of the MAKS is theoretically predefined, a theory-driven confirmatory approach was prioritised to evaluate structural equivalence before considering exploratory alternatives.

Confirmatory factor analysis was conducted using

maximum likelihood estimation to allow comparison with the original MAKS validation framework. However, given the clear evidence of non-normal item distributions, the use of maximum likelihood (ML) estimation may have resulted in conservative or unstable fit indices. As such, CFA findings are interpreted primarily as diagnostic indicators of structural viability rather than confirmatory tests of model adequacy. Factor loadings for individual items were inspected to determine their contribution to the respective constructs, with values above 0.40 regarded as satisfactory indicators [24, 27].

The adequacy of the model was judged through multiple fit indices, including the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Criteria for acceptable model fit required CFI and TLI values greater than 0.90, RMSEA below 0.08, and SRMR below 0.08 [27, 28]. Collectively, these indices provided a robust appraisal of how well the hypothesized two-factor structure aligned with the observed data, thereby reinforcing the scale's structural validity.

3. RESULTS

The study involved 203 respondents with a mean age of 26 years, with the majority (70.4%) being under 25 years old. Most participants were female (68.0%), and the predominant education level was a bachelor's degree (74.9%), with smaller proportions holding a master's degree (7.4%), diploma (6.9%), secondary school education (8.4%), DrPH (2.0%), or primary school education (0.5%). In terms of occupation, nearly three-quarters (72.9%) were students, followed by employed individuals (20.2%), unemployed participants (6.4%), and one retiree (0.5%). A large proportion of respondents (81.3%) reported that they currently know or have known someone with a mental health problem, indicating high familiarity with the topic (Table 1).

3.1. Descriptive Analysis

Table 2 presents descriptive analyses of the 12 items of the Malay version of the Mental Health Knowledge Schedule (MAKS-M), with mean scores ranging from 2.764 (Item 8, reversed "Stress [English ver.]") to 4.650 (Item 10, "Bipolar Disorder (manic-depression)[English ver.]"). Items from the knowledge subscale (Items 1-6, with Item 6 reversed) generally scored between 2.837 and 4.187, with skewness values mostly negative except for the reversed items. Items from the attitude subscale (Items 7-12, with Items 8 and 12 reversed) showed means from 2.764 to 4.650, with several items (*e.g.*, Items 7 "Depression [English ver.]", 9 "Schizophrenia [English ver.]", and 10 "Bipolar Disorder (manic-depression)[English ver.]") displaying strong negative skewness. Kurtosis values varied, with some high positive values (*e.g.*, Item 9 = 3.158), particularly for items with extreme skewness. The minimum and maximum scores covered the full range (1-5; Appendix C).

Table 1. Participants sociodemographic information.

-	-	Mean	Count	Column N %
Age (years)	-	26	-	-
Age group	Less 25 years old	-	143	70.4%
	25 years old and above	-	60	29.6%
Gender	Male	-	65	32.0%
	Female	-	138	68.0%
Education level	Primary school	-	1	0.5%
	Secondary school	-	17	8.4%
	Diploma	-	14	6.9%
	Bachelor's degree	-	152	74.9%
	Master's degree	-	15	7.4%
	DrPH	-	4	2.0%
Occupation status	Student	-	148	72.9%
	Employed	-	41	20.2%
	Unemployed	-	13	6.4%
	Retiree	-	1	0.5%
Known, or have known, someone with a mental health problem	Yes	-	165	81.3%
	No	-	38	18.7%

Table 2. Summary descriptive statistics for MAKS-M subscales.

Subscale	Number of Items	Mean (SD)	Observed Range	Skewness Range	Kurtosis Range
Knowledge (Items 1-6*)	6	3.68 (0.62)	2.17 - 4.60	-0.73 to +0.08	-0.90 to +0.64
Attitude (Items 7-12*)	6	3.93 (0.79)	1.17 - 4.93	-1.86 to +0.10	-1.11 to +3.16

Notes: * Items 6, 8, and 12 are reversed items included in the total subscale scores.

Full item-level descriptive statistics are presented in Appendix C.

Table 3. Summary of reliability indices for MAKS-M subscales.

Subscale	McDonald's ω	Cronbach's α	GLB	α (after removing reversed items)	Interpretation
Knowledge	0.396	0.262	0.480	0.500	Poor reliability
Attitude	0.444	0.236	0.646	0.566	Poor reliability

Notes: - GLB: Greatest Lower Bound.

-Removal of reversed items (Items 6, 8, and 12) improves internal consistency but remains below acceptable thresholds ($\alpha \geq 0.70$).

-Full reliability diagnostics are presented in Appendix D.

The normality tests for the MAKS-M subscales indicated that both the knowledge (sum of Items 1-6, with Item 6 reversed) and attitude (sum of Items 7-12, with Items 8 and 12 reversed) scores deviated significantly from normality. For the knowledge subscale, the Kolmogorov-Smirnov test ($D = 0.109$, $p < .001$) and the Shapiro-Wilk test ($W = 0.979$, $p = .004$) both yielded significant results, suggesting non-normality. Similarly, for the attitude subscale, the Kolmogorov-Smirnov test ($D = 0.135$, $p < .001$) and the Shapiro-Wilk test ($W = 0.959$, $p < .001$) confirmed a significant departure from normality (Table 2).

3.2. Reliability Analysis

The reliability analysis of the MAKS-M subscales showed poor internal consistency in both the knowledge

and attitude domains. For the knowledge subscale (Items 1-6, with Item 6 reversed), McDonald's ω was 0.396, Cronbach's α was 0.262, and the Greatest Lower Bound (GLB) was 0.480, with the 95% confidence intervals (Table 3). Notably, Item 6 (reversed) "Most people with mental health problems... [English ver.]" correlated negatively with the overall scale (Appendix D) [25].

Similarly, for the attitude subscale (Items 7-12, with Items 8 and 12 reversed), McDonald's ω was 0.444, Cronbach's α was 0.236, and the GLB was 0.646. Item 8 (reversed), "Stress [Eng ver.]", also showed a negative correlation with the scale.

The item-level reliability analysis for the MAKS-M knowledge subscale (Items 1-6, with Item 6 reversed) showed that most items had low item-rest correlations.

Item-rest correlations ranged from 0.046 (Item 5) to 0.413 (Item 4). Notably, Item 6 (reversed), “Most people with mental health problems... [English ver.]” had a negative item-rest correlation (-0.200). Cronbach’s α values if an item was deleted showed that removing Item 6 would substantially increase α from 0.262 to 0.500.

The item-level reliability analysis for the MAKS-M attitude subscale (Items 7-12, with Items 8 and 12 reversed) revealed notable variability in item performance. Most items showed moderate positive item-rest correlations (0.259 to 0.555), with Items 9, “Schizophrenia [Eng ver.]”, and 10, “Bipolar Disorder (manic-depression) [Eng ver.]”, having the highest correlations (0.524 and 0.555, respectively). In contrast, Item 8 (reversed), “Stress [Eng ver.]”, and Item 12 (reversed), “Grief [Eng ver.]”, had negative item-rest correlations (-0.316 and -0.145). Cronbach’s α values if an item was deleted showed that removing Item 8 would increase α substantially from 0.236 to 0.566, while removing Item 12 would also raise α to 0.439.

3.3. Construct Validity

The confirmatory factor analysis (CFA) for the MAKS-M yielded a significant chi-square result ($\chi^2(53) = 217.611, p < .001$), indicating some model misfit (Table 4). The fit indices showed suboptimal values, with CFI = 0.712, TLI = 0.641, and NFI = 0.659, all below the commonly accepted 0.90 threshold. RMSEA was 0.124 and SRMR was 0.109, both exceeding recommended cut-offs (< 0.08) (Appendix E).

From factor loading analysis of each items, revealed that most items had statistically significant loadings on

their respective factors ($p < .05$), except for Item 6 (reversed) “Most people with mental health problems... [English ver.]” in the knowledge subscale, and Items 8 (reversed) “Stress [Eng ver.] and 12 (reversed) “Grief [Eng ver.]” in the attitude subscale, which had low or negative loadings and non-significant p -values. Item 3 “Medication can be an effective treatment... [Eng ver.]”, 4 “Psychotherapy (e.g., talking therapy or counselling) can be an effective treatment... [Eng ver.]”, 9 “Schizophrenia [Eng ver.]”, and 10 “Bipolar Disorder (manic-depression) [Eng ver.]” showed the strongest loadings (> 0.50).

3.4. Convergent Validity

Table 5 presents the convergent validity analysis of the subscales, with correlations with BPIM. The Spearman’s correlation analysis indicated a significant positive relationship between the MAKS knowledge and MAKS attitude subscales ($\rho = 0.307, p < .001$). MAKS knowledge also showed a significant positive correlation with BIPM interest ($\rho = 0.230, p < .001$), and MAKS attitude correlated positively with BIPM interest as well ($\rho = 0.218, p = .002$). However, neither MAKS knowledge nor MAKS attitude correlated significantly with BIPM insight ($p > .05$).

4. DISCUSSION

4.1. Internal Consistency and Scale Reliability

The present findings should be interpreted not as a failed validation attempt, but as a methodological case study illustrating the psychometric vulnerabilities of direct scale transfer in cross-cultural mental health literacy measurement.

Table 4. Summary of confirmatory factor analysis (CFA) fit indices for MAKS-M.

Fit Index	Value	Interpretation
CFI	0.712	Below the acceptable threshold (< 0.90)
TLI	0.641	Below the acceptable threshold (< 0.90)
RMSEA	0.124	Poor fit; exceeds recommended cut-off (< 0.08)
SRMR	0.109	Poor fit; exceeds recommended cut-off (< 0.08)

Notes: - Fit indices indicate that the hypothesized two-factor structure of the MAKS is not supported in the Malaysian context.

- Complete CFA model parameters and factor loadings are provided in Appendix E.

Table 5. Convergent validity.

Variable	-	MAKS Knowledge	MAKS Attitude	BIPM Insight	BIPM Interest
1. MAKS knowledge	Spearman's rho	-			
	p -value	-			
2. MAKS attitude	Spearman's rho	0.307 ***	-		
	p -value	$< .001$	-		
3. BIPM insight	Spearman's rho	-0.073	-0.064	-	
	p -value	0.300	0.366	-	
4. BIPM interest	Spearman's rho	0.230 ***	0.218 **	0.089	-
	p -value	$< .001$	0.002	0.208	-

Note: * $p < .05$, ** $p < .01$, *** $p < .001$.

The low internal consistency observed in both subscales of the MAKs-M suggests deeper structural and contextual issues that extend beyond measurement error. One likely cause is the presence of reversed items, which are known to compromise reliability in cross-cultural adaptations, especially in multilingual contexts like Malaysia [12, 29]. The reversed items were retained during initial validation to maintain conceptual equivalence with the original MAKs and to empirically test whether previously reported item instability also emerged in the Malaysian context. Still, these items can create unintended cognitive load, where respondents-particularly those with varying literacy levels or second-language fluency-may misinterpret item direction, leading to inconsistent or contradictory responses. This semantic complexity may obscure the intended construct and undermine the coherence of the scale [12]. Additionally, the weak reliability indices observed may reflect not only the cognitive burden of reversed items in translated scales but also distributional characteristics; several reversed items demonstrated marked negative skew and peaked distributions, limiting variance and attenuating item-total correlations.

Another contributing factor may be the conceptual ambiguity within the subscales themselves. The knowledge subscale, for instance, includes items tapping both factual understanding and social perceptions of mental health (*e.g.*, recognizing schizophrenia vs. beliefs about treatment effectiveness), which may not form a unidimensional construct. Similarly, the attitude subscale includes items ranging from general acceptance of mental health conditions to more specific views about treatment and disclosure [4, 12].

These findings reinforce prior arguments that the MAKs may be better suited for item-level or domain-specific interpretation rather than aggregate scoring, particularly in non-Western contexts. A scale with unreliable subscales may produce misleading conclusions about the public's mental health knowledge and attitudes, particularly when used to evaluate interventions or track population-level changes. In research contexts, this limits the scale's utility in hypothesis testing and weakens its psychometric credibility. In policy or educational settings, it may lead to erroneous assumptions about the effectiveness of mental health literacy campaigns or the prevalence of stigma [30, 31].

4.2. Item-Level and Structural Validity

The item-level analysis of the MAKs-M revealed inconsistencies that raise concerns about the scale's structural validity. While some items showed acceptable alignment with their intended subscales, others-particularly the reversed items-performed poorly, with negative or near-zero item-rest correlations and non-significant factor loadings. These discrepancies suggest that certain items did not meaningfully contribute to the latent constructs they were intended to measure, pointing to potential issues in semantic clarity, conceptual alignment, or cultural relevance of specific statements [32].

One possible cause of weak item performance is the mismatch between item content and local interpretations of mental health. Items originally developed in Western contexts may not transfer seamlessly into the Malaysian cultural and linguistic environment, where beliefs, idioms, and stigma surrounding mental illness differ [33]. For example, terms like "treatment" or "disclosure" may carry different social connotations in collectivist cultures, where family reputation and social harmony play a stronger role in shaping attitudes. If the item content fails to reflect culturally embedded constructs of mental health knowledge and stigma, it may lack salience or be interpreted inconsistently by respondents [34].

Another contributing factor is item redundancy or ambiguity. Items that are too broadly worded may allow for varied interpretations, especially among respondents with differing educational backgrounds or mental health familiarity. Conversely, some items may be overly narrow or context-specific, limiting their generalizability and weakening their structural contribution. Moreover, the tendency of some items to elicit ceiling effects-where a large proportion of respondents choose the most favorable response-can reduce variance and inflate apparent item irrelevance, even if the content is meaningful [35].

Exploratory factor analysis (EFA) may be warranted to empirically reassess item groupings and potentially redefine subscale boundaries that better reflect local mental health literacy constructs. EFA was not initially prioritized because the primary objective was to test whether the original theoretical structure could be replicated in a new cultural context; only after structural non-equivalence was demonstrated does exploratory redevelopment become methodologically appropriate.

4.3. Model Fit and Factor Structure Comparison

The confirmatory factor analysis (CFA) of the MAKs-M revealed substantial misfit to the hypothesized two-factor structure, raising important concerns about the applicability of the original MAKs model in the Malaysian context [36]. In this context, CFA results are best understood as diagnostic evidence of structural misfit, highlighting the limits of direct construct transfer rather than as confirmatory tests of factorial validity.

While a two-factor structure-separating knowledge and attitude components-has been supported in some previous validations, these findings suggest that the theoretical distinction may not hold consistently across cultural settings [37]. Rather than viewing the poor model fit solely as model failure, these CFA results indicate that the original MAKs two-factor structure does not transfer adequately to the Malaysian context, suggesting a lack of structural equivalence.

Several factors may underlie the poor model fit. The inclusion of reversed items with weak or negative factor loadings (*e.g.*, Items 6, 8, and 12) likely distorted the factorial structure [38]. The predominance of negatively skewed distributions, particularly for reversed and high-endorsement items, likely reduced item variability and impaired inter-item associations. These challenges align

with prior reports that reversed items and ceiling effects commonly distort psychometric performance in cross-cultural mental health literacy instruments [38].

The relative homogeneity of the sample represents a design-linked psychometric constraint rather than merely a demographic limitation. The predominance of young, university-educated participants, combined with high prior exposure to mental health issues (81%), likely produced elevated endorsement of knowledge and attitude items. This resulted in pronounced ceiling effects and negative skewness, reducing item variance. Restricted variance attenuates inter-item correlations and covariance estimates, which in turn mechanically depress reliability coefficients and impair confirmatory factor analysis (CFA) model recovery. Thus, part of the observed psychometric instability may reflect limited score dispersion within this specific sample rather than purely structural flaws in the instrument.

The poor separation of subscales suggests that the MAK5-M may assess a hybrid construct that incorporates informational knowledge, personal beliefs, and perceived social norms. This raises the possibility that respondents are not reporting 'knowledge' in a positivist sense but rather articulating culturally mediated beliefs about mental health. Cultural norms and lived experiences may shape mental health beliefs in ways that blur the line between factual understanding and value-laden opinions [34]. For instance, endorsing a belief in the effectiveness of treatment may reflect both knowledge and optimism about recovery, which could reduce the clarity of the two-factor model.

Comparatively, previous adaptations of the MAK5 in Asian populations have encountered similar difficulties. Studies in Singapore and China reported challenges in maintaining the original factor structure, often necessitating model respecification or item removal to achieve acceptable fit [14, 15]. These patterns suggest that the MAK5 may require regional adaptation rather than direct translation, as its structural validity does not appear universally stable across contexts.

4.4. Cross Language Correlation

In this study, while full-scale back-translation procedures were followed, the results suggest that semantic and pragmatic differences between English and Malay may have compromised the integrity of certain items, particularly those with complex or nuanced phrasing.

The difficulties associated with reversed items in the MAK5-M likely reflect method effects rather than simple translation errors. Cross-cultural psychometric research has consistently demonstrated that negatively worded items may introduce systematic variance unrelated to the intended construct, often forming artificial method factors. In multilingual contexts, these effects are amplified by cognitive and linguistic asymmetry, as negation structures may alter sentence emphasis and processing demands [12, 28]. In Malay, explicit negative markers can increase syntactic complexity and pragmatic ambiguity, potentially triggering acquiescence or harmony-driven response

tendencies [28]. Thus, the instability observed is not a critique of reversed wording per se, but evidence of interaction between language structure, cognitive load, and culturally patterned response styles [39].

Another challenge in cross-language correlation lies in the cultural framing of mental health concepts. For instance, terms such as "mental illness," "treatment," or "disclosure" may evoke different connotations across languages and cultures [40]. Even if direct translation is linguistically accurate, the psychological and social meanings embedded in these words may vary, affecting how respondents relate to and interpret each item [41].

The implications for cross-cultural research are significant. If conceptual equivalence is not adequately preserved, any observed differences in scores between linguistic groups may reflect measurement artifacts rather than true differences in mental health literacy. This limits the utility of the MAK5-M in multicultural studies or in evaluating interventions across linguistically diverse populations [41].

4.5. Implications for Cross-Cultural Application of MAK5

The findings of this study underscore the challenges of applying mental health literacy instruments across cultural contexts without rigorous adaptation. Although the MAK5 was originally developed to assess universal aspects of mental health knowledge and stigma, its application in Malaysia revealed limitations that extend beyond linguistic translation.

In collectivist Malaysian settings, mental health understanding is influenced by family interdependence, moral attributions, and religious frameworks [34]. These factors reduce the distinction between what people believe to be true, what they feel socially expected to endorse, and what they personally understand. These cultural underpinnings influence not only the interpretation of mental health concepts but also the acceptability of certain attitudes and behaviours [42]. For example, questions about disclosure or treatment-seeking may invoke concerns about shame or social harmony, which are not explicitly addressed in the original MAK5. As such, items developed in Western settings may not resonate with or fully capture the lived experiences and stigma mechanisms present in non-Western societies [43].

4.6. Implications for Redevelopment

Rather than viewing these results solely as translation shortcomings, the findings highlight foundational measurement weaknesses within the MAK5 framework. The inability of the MAK5-M to demonstrate acceptable psychometric performance underlines the need to reconceptualise mental health knowledge assessment in Malaysia from the ground up. The evidence indicates that future work should prioritise the redevelopment of item content rather than incremental revision. Psychometric instability in multiple international adaptations suggests that the MAK5 construct architecture itself may lack robustness across cultural settings.

4.7. Contribution Beyond MAKS-M

Beyond evaluating the Malay version of the MAKS, this study offers broader methodological insights into cross-cultural mental health measurement. It demonstrates that reversed items are particularly vulnerable to linguistic and cultural distortion, often producing systematic psychometric failure rather than random noise. The findings highlight that mental health “knowledge” and “attitude” may not function as clearly separable constructs in non-Western contexts, where factual understanding, moral beliefs, and social norms are closely intertwined. By treating poor psychometric performance as informative rather than dismissible, this study illustrates the value of explicitly testing construct viability rather than assuming structural stability during scale adaptation.

5. STRENGTHS AND LIMITATIONS

One key strength lies in its comprehensive psychometric evaluation, employing a range of reliability indices (Cronbach’s α , McDonald’s ω , GLB), item-level diagnostics, and confirmatory factor analysis. This multi-method approach provided nuanced insights into the scale’s internal structure and performance. Additionally, the inclusion of convergent validity testing with the Balanced Index of Psychological Mindedness (BIPM) adds credibility to the interpretation of subscale associations, particularly with regard to the “interest” dimension of psychological mindedness.

The study also benefits from a relatively diverse participant pool in terms of educational and occupational background, although it was predominantly composed of young and university-educated individuals. The methodological rigor in translation-following standard forward- and back-translation procedures adds value by ensuring at least a foundational level of linguistic equivalence between the original and Malay versions.

However, several limitations must be acknowledged. First and foremost, the bilingual administration of instruments, while intended to support cross-language comparison, may have introduced priming effects that influenced item interpretation and response consistency.

The second limitation relates to the distributional characteristics of several MAKS-M items. Item-level analyses revealed pronounced negative skewness and ceiling effects, particularly for items assessing recognition of common mental health conditions (*e.g.*, depression, schizophrenia, bipolar disorder). Such skewed distributions reduce item variance and can attenuate inter-item correlations, thereby adversely affecting confirmatory factor analysis (CFA) results. As CFA using maximum likelihood estimation assumes approximate multivariate normality, deviations from normality may have contributed to the observed model misfit. These issues limit the current version’s suitability for research, policy evaluation, or community use without substantial revision or redevelopment.

Thirdly, the sample, while adequate in size, was not demographically representative of the broader Malaysian population. The high proportion of participants with prior

exposure to mental health issues (over 80%) and tertiary education may have introduced a bias toward more favorable attitudes and higher knowledge levels, thus contributing to ceiling effects and limiting generalizability.

Fourth, the reliance on self-report measures introduces the possibility of social desirability bias, particularly in items related to mental health stigma. Respondents may have selected more socially acceptable answers, especially in a survey format without interviewer clarification. This may have obscured true variability in attitudes and knowledge, further complicating reliability and validity interpretation.

Finally, while this study focused on psychometric properties, it did not incorporate qualitative feedback on item comprehension or cultural relevance. Such insights could have enriched the interpretation of problematic items and informed more targeted revisions.

CONCLUSION

This study demonstrated why the MAKS, in its current form, does not transfer psychometrically across cultures, thereby highlighting the need for redevelopment rather than incremental translation. These findings provide empirical guidance for researchers engaged in cross-cultural measurement, particularly regarding the risks posed by reversed items and poorly delineated constructs. The MAKS-M, in its current iteration, falls short of acceptable psychometric standards and should not be used for composite scoring in Malaysia in its current form.

RECOMMENDATIONS

Based on the findings of this study, it is recommended that the Malay version of the Mental Health Knowledge Schedule (MAKS-M) undergo redevelopment before being adopted for widespread use in Malaysian research or practice. The current psychometric limitations—particularly poor internal consistency, misfitting factor structure, and problematic reversed items—underscore the need for both linguistic and conceptual revisions to improve the tool’s validity and reliability.

One of the most urgent adjustments involves the removal or rewording of reversed items (Items 6, 8, and 12). These items consistently produced negative item-rest correlations and weak factor loadings, suggesting that their negative phrasing may have introduced confusion or cognitive strain for respondents [44]. Future Malay adaptations of mental health literacy tools should avoid negatively worded items and consider rephrasing these items into positively worded statements while maintaining their intended construct. For example, instead of negatively phrased items that require mental reversal (*e.g.*, “Most people with mental health problems go to prison”), a direct but positively framed alternative (*e.g.*, “People with mental health problems are not more likely to go to prison”) may improve interpretability without sacrificing meaning.

Additionally, several items may require conceptual clarification or cultural adaptation to better align with local beliefs and practices around mental health. Items that

reference treatment, disclosure, or stigma should be evaluated for cultural salience, perhaps by incorporating commonly understood terms or examples that reflect Malaysian mental health realities [45]. This could include references to family involvement in care decisions, religious interpretations of illness, or community-level stigma, which may not be captured in the original item phrasing.

Future validation designs should draw from more demographically diverse groups-including rural, older, and less-educated populations-to ensure adequate variability and to reduce ceiling effects. Additionally, qualitative methods such as cognitive interviews, focus groups, and expert panel reviews should be integrated into future validation phases. These approaches can help identify misunderstood items, ambiguous language, or culturally incongruent content. Exploratory factor analysis (EFA) is also recommended in the early stages of revision to empirically explore the underlying structure of the scale and to assess whether the original two-factor model remains appropriate, or whether an alternative configuration better reflects local constructs of mental health literacy.

Future research should treat the MAKS as a conceptual starting reference rather than a scale ready for adaptation. Item-level redevelopment informed by qualitative and cross-cultural methods may be required to produce a psychometrically stable measure of mental health knowledge. Any endeavor of MAKS validation should include more demographically diverse samples to enhance the generalizability of the findings. Validation in populations varying by age, education level, geographic location (including rural areas), and cultural background within Malaysia will be essential to ensure the scale functions reliably across subgroups.

AUTHORS' CONTRIBUTIONS

It is hereby acknowledged that all authors have accepted responsibility for the manuscript's content and consented to its submission. They have meticulously reviewed all results and unanimously approved the final version of the manuscript.

LIST OF ABBREVIATIONS

CFA = Confirmatory Factor Analysis

MAKS-M = Mental Health Knowledge Schedule

BIPM = Balanced Index of Psychological

GLB = Mindedness Greatest Lower Bound

RMSEA = Root Mean Square Error of Approximation

SRMR = Standardized Root Mean Square Residual

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Ethical approval for this study was obtained from the Research Ethics Committee, University Malaysia Sabah, Malaysia, with approval code JKetika 4/25(27).

HUMAN AND ANIMAL RIGHTS

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2013.

CONSENT FOR PUBLICATION

Written informed consent was obtained from all participants for their participation in the study, and all data were anonymized prior to analysis and reporting. Therefore, no additional consent for publication of individual data is required.

STANDARDS OF REPORTING

STROBE guidelines were followed.

AVAILABILITY OF DATA AND MATERIALS

The data supporting the findings of this article are available in the Zenodo repository at <https://doi.org/10.5281/zenodo.21126479>.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

APPENDIX

Appendix A: MAKS Malay Translation

Item	English MAKS	Malay Translation
	Mental Health Knowledge Schedule	Jadual Pengetahuan Mengenai Kesihatan Mental
1	Most people with mental health problems want to have paid employment.	Ramai orang yang ada gejala kesihatan mental ingin pekerjaan yang dibayar gaji.
2	If a friend had a mental health problem, I know what advice to give them to get professional help	Kalau kawan saya ada gejala kesihatan mental, saya tahu nasihat yang boleh diberikan untuk mereka mendapatkan bantuan pakar.
3	Medication can be an effective treatment for people with mental health problems	Mengambil ubat dari hospital berkesan untuk merawat gejala kesihatan mental.
4	Psychotherapy (eg talking therapy or counseling) can be an effective treatment for people with mental health problems	Psikoterapi (kaunseling dan terapi meluahkan perasaan) berkesan untuk merawat gejala kesihatan mental.
5	People with severe mental health problems can fully recover	Mereka yang mempunyai gejala kesihatan mental yang teruk boleh sembuh sepenuhnya.

Appendix A contd.....

Item	English MAKS	Malay Translation
6	Most people with mental health problems go to a healthcare professional to get help	Ramai orang yang ada gejala kesihatan mental dapatkan bantuan dari pakar perubatan. (reversed)
7	Depression	Kemurungan
8	Stress	Stres (reversed)
9	Schizophrenia	Skizofrenia
10	Bipolar Disorder (manic-depression)	Kecelaruhan Bipolar (episod mania - kemurungan)
11	Drug addiction	Ketagihan Dadah
12	Grief	Kehibaan (Perasaan kehilangan sesuatu yang dikasihi) (reversed)

Appendix B: MAKS-Malay

Arahan: Bagi pernyataan 1 - 6, sila tandakan pada satu petak sahaja. Gejala kesihatan mental merujuk kepada keadaan-keadaan yang sepatutnya dirawat oleh ahli-ahli penjagaan kesihatan.							
		Sangat setuju	Agak setuju	Neutral	Kurang setuju	Sangat Tidak Setuju	Tidak Tahu
1	Ramai orang yang ada gejala kesihatan mental ingin pekerjaan yang dibayar gaji.						
2	Kalau kawan saya ada gejala kesihatan mental, saya tahu nasihat yang boleh diberikan untuk mereka mendapatkan bantuan pakar.						
3	Mengambil ubat dari hospital berkesan untuk merawat gejala kesihatan mental.						
4	Psikoterapi (kaunseling dan terapi meluahkan perasaan) berkesan untuk merawat gejala kesihatan mental.						
5	Mereka yang mempunyai gejala kesihatan mental yang teruk boleh sembuh sepenuhnya.						
6	Ramai orang yang ada gejala kesihatan mental dapatkan bantuan dari pakar perubatan.						
Arahan: Nyatakan sama ada kondisi berikut adalah jenis penyakit mental dengan menandakan satu kotak sahaja.							
7	Kemurungan						
8	Stres						
9	Schizophrenia						
10	Kecelaruhan Bipolar (episod mania - kemurungan)						
11	Ketagihan Dadah						
12	Kehibaan (Perasaan kehilangan sesuatu yang dikasihi)						

Appendix C: Full Item Descriptive Statistics of MAKS-M

Item	Mean	Std. Deviation	Skewness	Std. Error of Skewness	Kurtosis	Std. Error of Kurtosis	Minimum	Maximum
1. Ramai orang yang ada gejala kesihatan mental ingin pekerjaan yang dibayar gaji.	3.985	0.876	-0.730	0.171	0.638	0.340	1.000	5.000
2. Kalau kawan saya ada gejala kesihatan mental, saya tahu nasihat yang boleh diberikan untuk mereka mendapatkan bantuan pakar.	3.833	0.809	-0.534	0.171	0.303	0.340	1.000	5.000
3. Mengambil ubat dari hospital berkesan untuk merawat gejala kesihatan mental.	3.897	0.887	-0.656	0.171	0.460	0.340	1.000	5.000
4. Psikoterapi (kaunseling dan terapi meluahkan perasaan) berkesan untuk merawat gejala kesihatan mental.	4.187	0.774	-0.661	0.171	-0.081	0.340	2.000	5.000
5. Mereka yang mempunyai gejala kesihatan mental yang teruk boleh sembuh sepenuhnya.	3.310	0.958	0.024	0.171	-0.464	0.340	1.000	5.000
6. Ramai orang yang ada gejala kesihatan mental dapatkan bantuan dari pakar perubatan. (reversed)	2.837	1.103	0.081	0.171	-0.900	0.340	1.000	5.000
7. Kemurungan	4.527	0.692	-1.504	0.171	2.174	0.340	2.000	5.000
8. Stres (reversed)	2.764	1.260	0.095	0.171	-1.110	0.340	1.000	5.000
9. Skizofrenia	4.586	0.755	-1.863	0.171	3.158	0.340	1.000	5.000

Appendix C contd.....

Item	Mean	Std. Deviation	Skewness	Std. Error of Skewness	Kurtosis	Std. Error of Kurtosis	Minimum	Maximum
10. Kecelaruhan Bipolar (episod mania - kemurungan)	4.650	0.638	-1.852	0.171	3.033	0.340	2.000	5.000
11. Ketagihan Dadah	3.616	1.255	-0.440	0.171	-1.023	0.340	1.000	5.000
12. Kehibaan (Perasaan kehilangan sesuatu yang dikasihi) (reversed)	3.463	1.263	-0.362	0.171	-0.987	0.340	1.000	5.000

Appendix D: Full item-level reliability tables**Scale Reliability Statistics: Knowledge subscale.**

Estimate	McDonald's ω	Cronbach's α	Greatest Lower Bound
Point estimate	0.396	0.262	0.480
95% CI lower bound	0.284	0.089	0.408
95% CI upper bound	0.509	0.409	0.606

Note: The following item correlated negatively with the scale: Item6_reversed.**Scale Reliability Statistics: Attitude subscale.**

Estimate	McDonald's ω	Cronbach's α	Greatest Lower Bound
Point estimate	0.444	0.236	0.646
95% CI lower bound	0.347	0.051	0.534
95% CI upper bound	0.541	0.391	0.742

Note: The following item correlated negatively with the scale: Item8_reversed.**Individual Item Reliability Statistics: Knowledge subscale.**

	If item dropped	
Item	Cronbach's α	Item-rest correlation
Item1	0.175	0.183
Item2	0.210	0.141
Item3	0.080	0.295
Item4	0.013	0.413
Item5	0.284	0.046
Item6_reversed	0.500	-0.200

Individual Item Reliability Statistics: Attitude subscale.

	If item dropped	
Item	Cronbach's α	Item-rest correlation
Item7	0.026	0.419
Item8_reversed	0.566	-0.316
Item9	-0.072	0.524
Item10	-0.036	0.555
Item11	0.027	0.259
Item12_reversed	0.439	-0.145

Appendix E: Full CFA factor loading table with fit indices

Confirmatory Factor Analysis (CFA) - Chi-square test.

Model	X ²	df	p
Baseline model	637.323	66	
Factor model	217.611	53	< .001

Note: The estimator is ML.

Fit indices.

Index	Value
Comparative Fit Index (CFI)	0.712
Tucker-Lewis Index (TLI)	0.641
Bentler-Bonett Normed Fit Index (NFI)	0.659
Root mean square error of approximation (RMSEA)	0.124
Standardized root mean square residual (SRMR)	0.109
Goodness of fit index (GFI)	0.991

Factor Loadings.

						95% Confidence Interval	
Factor	Indicator	Estimate	Std. Error	z-value	p	Lower	Upper
Knowledge	Item1	0.296	0.076	3.888	< .001	0.147	0.446
	Item2	0.265	0.071	3.718	< .001	0.125	0.404
	Item3	0.531	0.085	6.273	< .001	0.365	0.696
	Item4	0.504	0.076	6.648	< .001	0.355	0.652
	Item5	0.209	0.092	2.277	0.023	0.029	0.388
	Item6_reversed	-0.119	0.099	-1.202	0.229	-0.312	0.075
Attitude	Item7	0.409	0.046	8.908	< .001	0.319	0.498
	Item8_reversed	0.130	0.093	1.400	0.162	-0.052	0.311
	Item9	0.690	0.044	15.692	< .001	0.604	0.776
	Item10	0.575	0.037	15.399	< .001	0.501	0.648
	Item11	0.306	0.091	3.358	< .001	0.127	0.484
	Item12_reversed	-0.080	0.093	-0.865	0.387	-0.262	0.102

REFERENCES

[1] Jorm AF. Mental health literacy. Public knowledge and beliefs about mental disorders. *Br J Psychiatry* 2000; 177(5): 396-401. <http://dx.doi.org/10.1192/bjp.177.5.396> PMID: 11059991

[2] Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. "Mental health literacy": A survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. *Med J Aust* 1997; 166(4): 182-6. <http://dx.doi.org/10.5694/j.1326-5377.1997.tb140071.x> PMID: 9066546

[3] Sodi T, Quarshie ENB, Oppong Asante K, et al. Mental health literacy of school-going adolescents in sub-Saharan Africa: A regional systematic review protocol. *BMJ Open* 2022; 12(9): e063687. <http://dx.doi.org/10.1136/bmjopen-2022-063687> PMID: 36127093

[4] Evans-Lacko S, Little K, Meltzer H, et al. Development and psychometric properties of the mental health knowledge schedule. *Can J Psychiatry* 2010; 55(7): 440-8. <http://dx.doi.org/10.1177/070674371005500707> PMID: 20704771

[5] Miteva D, Georgiadis F, McBroom L, et al. Impact of language proficiency on mental health service use, treatment and outcomes: "Lost in Translation". *Compr Psychiatry* 2022; 114: 152299. <http://dx.doi.org/10.1016/j.comppsy.2022.152299> PMID: 35220037

[6] Thong VC, Pang NTP, Kamu A, Hadi NMN, Lasimbang HB, Ho CM. Validation Study of Malay (Bahasa Melayu) Version of the Dialectical Behavior Therapy - Ways of Coping Checklist (DBT-WCCL) and its Psychometric Properties. *Open Psychol J* 2025; 18(1): e18743501400689. <http://dx.doi.org/10.2174/0118743501400689250617113425>

[7] Bhugra D, Watson C, Wijesuriya R. Culture and mental illnesses. *Int Rev Psychiatry* 2021; 33(1-2): 1-2. <http://dx.doi.org/10.1080/09540261.2020.1777748> PMID: 33599563

[8] Chan WX, Salmah BM. Cultural beliefs, mental health literacy, and help-seeking behaviours among young sarawak indigenous adults. *J Cogn Sci Human Dev* 2023; 9(2): 209-28. <http://dx.doi.org/10.33736/jcshd.5930.2023>

[9] Haslam N, Tse JSY. Public awareness of mental illness: Mental health literacy or concept creep? *Australas Psychiatry* 2025; 33(1): 18-20. <http://dx.doi.org/10.1177/10398562241292202> PMID: 39402888

[10] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures.

- Spine 2000; 25(24): 3186-91.
<http://dx.doi.org/10.1097/00007632-200012150-00014> PMID: 11124735
- [11] Abdullah A, Liew SM, Salim HS, Ng CJ, Chinna K. Health literacy research in Malaysia: A scoping review. *Sains Malays* 2020; 49(5): 1021-36.
<http://dx.doi.org/10.17576/jsm-2020-4905-07>
 - [12] Munawar K, Mukhtar F, Choudhry FR, Ng ALO. Mental health literacy: A systematic review of knowledge and beliefs about mental disorders in Malaysia. *Asia-Pac Psychiatry* 2022; 14(1): e12475.
<http://dx.doi.org/10.1111/appy.12475> PMID: 33963816
 - [13] Wei Y, McGrath PJ, Hayden J, Kutcher S. Measurement properties of mental health literacy tools measuring help-seeking: A systematic review. *J Ment Health* 2017; 26(6): 543-55.
<http://dx.doi.org/10.1080/09638237.2016.1276532> PMID: 28355928
 - [14] He XY, Tan WY, Guo LL, Ji YY, Jia FJ, Wang SB. Mental health literacy among urban and rural residents of Guangdong Province, China. *Risk Manag Healthc Policy* 2024; 17: 2305-18.
<http://dx.doi.org/10.2147/RMHP.S479868> PMID: 39371938
 - [15] Lee JE, Goh ML, Yeo SF. Mental health awareness of secondary schools students: Mediating roles of knowledge on mental health, knowledge on professional help, and attitude towards mental health. *Heliyon* 2023; 9(3): e14512.
<http://dx.doi.org/10.1016/j.heliyon.2023.e14512> PMID: 36950622
 - [16] Bagubair AA, Dapari R. Factors associated with mental health literacy among healthcare workers in a university hospital in Selangor. *Clin Epidemiol Glob Health* 2025; 35: 102142.
<http://dx.doi.org/10.1016/j.CEGH.2025.102142>
 - [17] Singh S, Zaki RA, Farid NDN, Kaur K. The determinants of mental health literacy among young adolescents in Malaysia. *Int J Environ Res Public Health* 2022; 19(6): 3242.
<http://dx.doi.org/10.3390/ijerph19063242> PMID: 35328930
 - [18] Sperber AD. Translation and validation of study instruments for cross-cultural research. *Gastroenterology* 2004; 126(Suppl. 1): S124-8.
<http://dx.doi.org/10.1053/j.gastro.2003.10.016> PMID: 14978648
 - [19] Pheh KS, Ong WHA, Low SK, Tan CS, Kok JK. The Malay version of the mental health knowledge schedule: A preliminary study. *Malays J Psychiatry* 2017; 26(1): 10-4.
 - [20] Hambleton RK. The next generation of the ITC test translation and adaptation guidelines. *Eur J Psychol Assess* 2001; 17(3): 164-72.
<http://dx.doi.org/10.1027//1015-5759.17.3.164>
 - [21] Borg DN, Bach AJE, O'Brien JL, Sainani KL. Calculating sample size for reliability studies. *PM R* 2022; 14(8): 1018-25.
<http://dx.doi.org/10.1002/pmrj.12850> PMID: 35596122
 - [22] Gunawan J, Marzilli C, Aunguroch Y. Establishing appropriate sample size for developing and validating a questionnaire in nursing research. *Belitung Nurs J* 2021; 7(5): 356-60.
<http://dx.doi.org/10.33546/bnj.1927> PMID: 37496511
 - [23] Amiruddin M, Kassim M, Pang NTP, *et al.* Validation of Bahasa Malaysia version of psychological mindedness in a university population. *IIUM Med J Malaysia* 2021; 20(2): 81-8.
<http://dx.doi.org/10.31436/imjm.v20i2.1718>
 - [24] Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011; 2: 53-5.
<http://dx.doi.org/10.5116/ijme.4dfb.8dfd> PMID: 28029643
 - [25] Tavakol M, Wetzel A. Factor Analysis: A means for theory and instrument development in support of construct validity. *Int J Med Educ* 2020; 11: 245-7.
<http://dx.doi.org/10.5116/ijme.5f96.0f4a> PMID: 33170146
 - [26] Shan Y, Ji M. Cultural acceptability of mental health scales. *New Front Transl Stud* 2024; 37-39: 37-9.
http://dx.doi.org/10.1007/978-981-97-1727-9_5
 - [27] Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. Cengage Learning 2019.
 - [28] Lee S T, van Heuven W J B, Price J M, Leong C X R. Translation norms for Malay and English words: The effects of word class, semantic variability, lexical characteristics, and language proficiency on translation. *Behav Res* 2023; 55: 3585-601.
<http://dx.doi.org/10.3758/s13428-022-01977-3>
 - [29] Helms JE, Henze KT, Sass TL, Mifsud VA. Treating Cronbach's Alpha Reliability coefficients as data in counseling research. *Couns Psychol* 2006; 34(5): 630-60.
<http://dx.doi.org/10.1177/0011000006288308>
 - [30] Xia Y, Yang Y. RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behav Res Methods* 2019; 51(1): 409-28.
<http://dx.doi.org/10.3758/s13428-018-1055-2> PMID: 29869222
 - [31] Edman JL, Koon TY. Mental illness beliefs in Malaysia: Ethnic and intergenerational comparisons. *Int J Soc Psychiatry* 2000; 46(2): 101-9.
<http://dx.doi.org/10.1177/002076400004600203> PMID: 10950358
 - [32] Li W, Reavley N. Recognition and beliefs about treatment for mental disorders in mainland China: A systematic review and meta-analysis. *Soc Psychiatry Psychiatr Epidemiol* 2020; 55(2): 129-49.
<http://dx.doi.org/10.1007/s00127-019-01799-3> PMID: 31641829
 - [33] Khan T, Hassali M, Tahir H, Khan A. A pilot study evaluating the stigma and public perception about the causes of depression and schizophrenia. *Iran J Public Health* 2011; 40(1): 50-6.
 PMID: 23113054
 - [34] Cheung GW, Cooper-Thomas HD, Lau RS, Wang LC. Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pac J Manage* 2024; 41(2): 745-83.
<http://dx.doi.org/10.1007/s10490-023-09871-y>
 - [35] Tan MM, Su TT, Ting RSK, Allotey P, Reidpath D. Religion and mental health among older adults: Ethnic differences in Malaysia. *Aging Ment Health* 2021; 25(11): 2116-23.
<http://dx.doi.org/10.1080/13607863.2020.1799939> PMID: 32741203
 - [36] Subu MA, Holmes D, Arumugam A, *et al.* Traditional, religious, and cultural perspectives on mental illness: A qualitative study on causal beliefs and treatment use. *Int J Qual Stud Health Well-being* 2022; 17(1): 2123090.
<http://dx.doi.org/10.1080/17482631.2022.2123090> PMID: 36097886
 - [37] Marquine MJ, Jimenez D. Cultural and linguistic proficiency in mental health care: A crucial aspect of professional competence. *Int Psychogeriatr* 2020; 32(1): 1-3.
<http://dx.doi.org/10.1017/S1041610219000541> PMID: 32008601
 - [38] Kyriazos T, Poga M, Kyriazos T, Poga M. Dealing with multicollinearity in factor analysis: The problem, detections, and solutions. *Open J Stat* 2023; 13(3): 404-24.
<http://dx.doi.org/10.4236/ojs.2023.133020>
 - [39] Montoya AK, Edwards MC. The poor fit of model fit for selecting number of factors in exploratory factor analysis for scale evaluation. *Educ Psychol Meas* 2021; 81(3): 413-40.
<http://dx.doi.org/10.1177/0013164420942899> PMID: 33994558
 - [40] Mubbashar MH, Farooq S. Mental health literacy in developing countries. *Br J Psychiatry* 2001; 179(1): 75.
<http://dx.doi.org/10.1192/bjp.179.1.75-a> PMID: 11435276
 - [41] Cruchinho P, López-Franco MD, Capelas ML, *et al.* Translation, cross-cultural adaptation, and validation of measurement instruments: A practical guideline for novice researchers. *J Multidiscip Healthc* 2024; 17: 2701-28.
<http://dx.doi.org/10.2147/JMDH.S419714> PMID: 38840704
 - [42] Virdiyanti R. Mental health dynamics in the context of collectivist culture: A study of indigenous communities in Indonesia. *J Ilmu Psikologi dan Kesehatan* 2025; 3(3): 99-112.
<http://dx.doi.org/10.47353/sikontan.v3i3.2779>
 - [43] Volkow ND, Gordon JA, Koob GF. Choosing appropriate language to reduce the stigma around mental illness and substance use disorders. *Neuropsychopharmacology* 2021; 46(13): 2230-2.
<http://dx.doi.org/10.1038/s41386-021-01069-4> PMID: 34276051
 - [44] Suárez-Álvarez J, Pedrosa I, Lozano L, García-Cueto E, Cuesta M,

Muñiz J. Using reversed items in Likert scales: A questionable practice. *Psicothema* 2018; 2(30): 149-58.
<http://dx.doi.org/10.7334/psicothema2018.33> PMID: 29694314

[45] Chong ST, Mohamad MS, Er AC. The mental health development in malaysia: History, current issue and future development. *Asian Soc Sci* 2013; 9(6): 1.
<http://dx.doi.org/10.5539/ass.v9n6p1>