Stigma and Health

An Evaluation of the Opening Minds Scale for Health Care Providers

Esario IV Daguman and Joanne Taylor

Online First Publication, March 3, 2025. https://dx.doi.org/10.1037/sah0000628

AMERICAN PSYCHOLOGICAL ASSOCIATION

Stigma and Health

ISSN: 2376-6972

https://doi.org/10.1037/sah0000628

An Evaluation of the Opening Minds Scale for Health Care Providers

Esario IV Daguman[1] and Joanne Taylor[2]
[1] Faculty of Health, Southern Cross University
[2] School of Psychology, Massey University

Health workers hold stigmatizing attitudes toward people with mental distress, and contact-based interventions have been developed to address these attitudes. However, measures used to evaluate interventions have mixed validity support, including measures developed with service user involvement. The present study intended to provide a psychometric examination of one such measure, the 15-item Opening Minds Scale for Health Care Providers (Kassam et al., 2012; Modgill et al., 2014). With 286 health worker participants from a paid academic survey platform, including 19 retest samples, the focus was on addressing methodological limitations of previous psychometric evaluations. Current data demonstrated mixed fit to the known correlated three-factor structure (standardized root-mean-square residual [SRMR] = .057, root-mean-square error of approximation [RMSEA] = .054, comparative fit index = .754, Tucker–Lewis fit index = .897) and poor fit with the unidimensional model (SRMR = .080, RMSEA = .084, comparative fit index = .418, Tucker–Lewis fit index = .755). Data fitting to the bifactor solution (a structure comprising a general factor and three lower order factors) was attempted but did not converge. Scores had strong internal consistency (ω_t = .73–.86), very weak test–retest reliability (r = −.46 to .21), and weak to moderate albeit statistically significant support for the interrelationship between the factors (r_s = .32–.55). There is some evidence to consider the 15-item Opening Minds Scale for Health Care Providers as a viable measure of stigmatizing attitudes. However, further robust and transparent evaluations are still needed to surface better validity support.

Clinical Impact Statement

The Opening Minds Scale for Health Care Providers scores in the study showed some support for the measure's known factor structure, high internal consistency, and low test–retest reliability, making it potentially useful for measuring short-term changes in stigma attitudes, such as the immediate effects of antistigma interventions. However, its potential measurement of an additional factor and inability to track long-term attitude changes suggest a need to refine the Opening Minds Scale for Health Care Providers or develop additional tools to capture temporary and stable stigma attitudes. Future research needs to consider improving the measure's validity and reliability and adapting it for use in various cultural contexts.

Keywords: stigma, attitudes, Opening Minds Scale for Health Care Providers, mental health, validation

Supplemental materials: https://doi.org/10.1037/sah0000628.supp

Georg Schomerus served as action editor.

Esario IV Daguman https://orcid.org/0009-0009-3319-7077
Joanne Taylor https://orcid.org/0000-0001-8502-3452

The study's methodology was considered in advance and has been made publicly available on the Open Science Framework and can be accessed at https://osf.io/tuh49/. The de-identified data and statistical analysis codes are also available in the same online repository.

The authors have no relevant financial or nonfinancial interests to disclose. This work was supported by the Post-Graduate Research Fund in the School of Psychology at Massey University. The authors acknowledge all the health worker survey participants of the study.

Our positionality statement follows: Because the performance and translation of research are influenced by the lens that shapes a person's understanding of the world, the authors briefly mention their positionality. Esario IV Daguman (he/him), a male Southeast Asian with experience in acute mental health crisis support, values the perspectives of "experts by experience" and a multiple-model view of health. His psychology education has socialized him into an open and thoughtful approach to research. Joanne Taylor (she/her) is a female Pākehā (European) clinical psychologist who teaches and trains clinical psychologists at a university. She is described by lived experience colleagues as an ally in supporting service user research and teaching.

Esario IV Daguman played a lead role in investigation, methodology, formal analysis, visualization, and writing–original draft and an equal role in conceptualization. Joanne Taylor played an equal role in conceptualization, resources, supervision, and writing–review and editing.

Correspondence concerning this article should be addressed to Esario IV Daguman, Faculty of Health, Southern Cross University, Coffs Harbour, NSW 2450, Australia. Email: e.daguman.10@student.scu.edu.au

Stigma toward people experiencing mental distress[1] remains a persistent global challenge, despite public health efforts (Ruesch, 2022). It is a multifaceted social process that begins with them-and-us attitudes and evolves into othering behaviors that give primacy to differences, rather than common grounds (Corrigan, 2005). Stigma can have profound and enduring negative impacts on individuals, including restricted medical treatment and support (Thornicroft et al., 2022). Unfortunately, it is established that many traditional workers within health services harbor stigmatizing attitudes (Livingston, 2020), which can be even more pronounced than those held by mental health workers (Gupta et al., 2024).

Programs that challenge stigmatizing attitudes among health workers exist, although their evaluation methodologies have been limited (Heim et al., 2020). Standardized stigma measures have not been used in many of these evaluations, and their psychometric properties have often been underreported or unavailable (Na et al., 2022). Such measurement practice does not help identify the strength and quality of knowledge bases linking programs to stigma reductions. It has been said that counteracting stigma globally requires investing "not just in the delivery of programmes, but also in their development and assessment" (Watts, 2022, p. 1397).

The Opening Minds Scale for Health Care Providers

Real investment in evaluating stigma measures is exemplified in the work on emerging the Opening Minds Scale for Health Care Providers (OMS-HC; Kassam et al., 2012). Developed by the "Opening Minds" of the Mental Health Commission of Canada, the OMS-HC is one of the most widely used self-report measures of stigmatizing attitudes toward mental distress worldwide and has strong psychometric evidence (Sastre-Rus et al., 2019). It is based on several frames of reference on mental health stigma (Corrigan, 2005; Link & Phelan, 2001; Thornicroft, 2006).

Development and Initial Evaluation

The OMS-HC instrument was developed through a multiphase process involving the adaptation of existing scales, consultation with stigma intervention experts and individuals with lived experience, and qualitative research with diverse health workers (Kassam et al., 2012). Integral to the development was the continuous involvement of individuals with lived experience of mental distress, ensuring the measure's comprehensiveness and appropriateness. The initial 20-item measure, piloted with a Canadian sample of health professionals, was subsequently refined through item analysis, resulting in a 12-item scale with two primary factors: *Attitudes of health care providers toward people with mental illness* and *Attitudes of health care providers toward disclosure of mental illness*.

Recognizing the critical role of social distance as a behavioral intention indicator, the OMS-HC authors regrouped and refined its factor structure using a larger, more diverse health worker sample (Modgill et al., 2014). This process led to a 15-item version with a robust three-factor structure encompassing *Attitudes*, *Disclosure and Help-seeking*, and *Social Distance*. Items such as "I would be reluctant to seek help if I had a mental illness" and "There is little I can do to help people with mental illness" exemplify the scale's content.

Succeeding Evaluations

Table 1 summarizes existing research on the factorial validity, internal consistency, and test–retest reliability of the English and translated versions of the OMS-HC. Four studies claimed they have evidence for the three-factor structure (Sapag et al., 2019; Valdivia Ramos et al., 2023; van der Maas et al., 2018; Zuaboni et al., 2021). However, some confirmed the structure only after excluding one (Chang et al., 2017) or two items (Happell et al., 2019), whereas others had different items mapping to the factors (Destrebecq et al., 2018; Zuaboni et al., 2021).

Several models have been proposed through various validation studies. Two studies suggested an alternative bifactor model, which includes a general stigma factor and correlated three lower order factors, though it lacked model-based reliability (Őri et al., 2020, 2023). Building on the bifactor solution, another study added an item and introduced a fourth factor named *Negative Views* (Carrara et al., 2023). By contrast, a consistent finding across several evaluations is that the scores from the OMS-HC do not fit with a unidimensional model (Carrara et al., 2023; Happell et al., 2019; Őri et al., 2020, 2023).

Scores from the measure demonstrated strong internal consistency overall ($\alpha = .69$–$.98$); nonetheless, it is more varied for the subscale scores, ranging from moderate to high ($\alpha = .48$–$.98$). On the other hand, the assessments of stability over time for the overall ($r = .99$; intraclass correlation coefficient = .95) and subscale scores (intraclass correlation coefficient = .84–.9) demonstrated high levels of reliability.

To a More Open and Thoughtful Evaluation

A diverse range of methodological approaches have been applied to evaluate the OMS-HC. However, a consistent lack of justification for these choices is evident. For example, many studies have employed confirmatory methods without providing the apparent reason that the structure is already known and may need not be explored (Carrara et al., 2023; Happell et al., 2019; Sapag et al., 2019; Valdivia Ramos et al., 2023; van der Maas et al., 2018). It is also impossible to assess potential bias arising from selective reporting of outcomes and analyses within these confirmatory studies given that they were without prespecified protocols that are accessible online.

There is another prevalent measurement practice that combines confirmatory and exploratory analyses without explicitly stating which data sets were used in each analysis (Chang et al., 2017; Őri et al., 2020, 2023; Zuaboni et al., 2021), raising potential overfitting issues in their measurement models. More importantly, there was a lack of clarity regarding the interpretation of model fit. Specifically, studies often failed to explicitly explain how they determined the adequacy of a model's fit to the data or the criteria used to establish model validity. A study relied on all fit indices exceeding established cutoffs (van der Maas et al., 2018), while others required only one

[1] The authors used the term "mental distress" to describe the psychological and emotional dimension of a complex experience without reference to specific diagnostic or illness classifications. This choice aligns with the shift away from pathologizing language (British Psychological Society, 2015) and reflects the preference of individuals with lived experience of distress (O'Hagan, 2018). The term "mental illness" is retained only when it appears in the items or factor names of the measure under study.

Table 1
Summary of Psychometric Properties of the OMS-HC

Researcher, year of publication, and country	Research population	*N*	Result		
			Factor structure	Internal consistency	Test–retest reliability
Kassam et al. (2012; Canada)	Health workers and health workers in training across different professional organizations and academic institutions	787	PCA: two factors, with 12 items (45% of the variance)	Total of 12 items: α = .78, Factor 1: α = .75, Factor 2: α = .72	Total: ICC = .66 (*n* = 112)
Modgill et al. (2014; Canada)	Health workers and students across health professions	1,523	EFA: three factors, with 15 items (45.3% of the variance)	Total: α = .79, Factor 1: α = .68, Factor 2: α = .67, Factor 3: α = .68	
Chang et al. (2017; Singapore)	Health workers in training, from medicine and nursing	1,002	CFA: three factors, with 15 items RMSEA = .096, CFI = .838, TLI = .804 EFA, ESEM: three factors, with 14 items (less Item 1) RMSEA = .069, CFI = .948, TLI = .909	14 items (less Item 1) Total: α = .75, Factor 1: α = .74, Factor 2: α = .60, Factor 3: α = .53	
Destrebecq et al. (2018; Italy)	Health workers in training, from nursing, physical therapy, dietetics, and occupational therapy	561	EFA: three factors, with 15 items—Items 1, 12, 13, 14, 18, and 20 loaded unto Factor 3 rather than Factor 1; Items 8, 9, 17, and 19 loaded unto Factor 1 rather than Factor 3 EFA: two factors, with 12 items—items for Factors 1 and 2 loaded interchangeably, contrary to the original mapping	15 items Factor 1: α = .76, Factor 2: α = .83, Factor 3: α = .82 12 items Factor 1: α = .74, Factor 2: α = .86	Total: *r* = .99 (over 1 week; *n* = 25)
van der Maas et al. (2018; Canada)	Health workers in community health settings	190	CFA: three factors, with 15 items SRMR = .07, RMSEA = .013, CFI = .996, TLI = .996	Total: α = .77, Factor 1: α = .79, Factor 2: α = .67, Factor 3: α = .72	
Happell et al. (2019; Australia, Finland, Iceland, Ireland, the Netherlands, and Norway)	Health workers in training, from nursing	423	CFA: one factor, with 15 items RMSEA = .11, WRMR = 1.7, CFI = .79, TLI = .75 CFA: three factors, with 15 items RMSEA = .07, WRMR = 1.17, CFI = .91, TLI = .89 CFA: three factors, with 13 items (no Items 6 and 18) RMSEA = .06, WRMR = 0.98, CFI = .95, TLI = .93	13 items (less Item 6 and 18) Factor 1: PSI = .61, Factor 2: PSI = .50, Factor 3: PSI = .69	

(*table continues*)

Table 1 (*continued*)

Researcher, year of publication, and country	Research population	*N*	Result		
			Factor structure	Internal consistency	Test–retest reliability
			Rasch analysis: three factors, with 13 items (no Items 6 and 18) Factor 1: $\chi^2(30) = 54.47$, $p = .004$, Item fit residual $SD = 1.21$, Person fit residual $SD = 1.05$ Factor 2: $\chi^2(12) = 12.42$, $p = .412$, Item fit residual $SD = 1.62$, Person fit residual $SD = 1.25$ Factor 3: $\chi^2(30) = 41.06$, $p = .086$, Item fit residual $SD = 0.52$, Person fit residual $SD = 1.05$, Unidimensionality test = 4.02%		
Sapag et al. (2019; Chile)	Health workers across health professions in public primary health care settings	803	SEM: three factors, 15 items SRMR = .048, RMSEA = .052, CFI = .832, TLI = .798	Total: α = .69, Factor 1: α = .53, Factor 2: α = .48, Factor 3: α = .60	
Őri et al. (2020; Hungary)	Health workers and health workers in training, from adult and child psychiatry	211	CFA: original 15 items RMSEA = .048, CFI = .818, TLI = .78 CFA: one factor, with 15 items RMSEA = .066, CFI = .642, TLI = .583 CFA: three factors based on EFA, with 15 items RMSEA = .045, CFI = .844, TLI = .812 CFA: three factors based on EFA, with 14 items (less Item 14) RMSEA = .043, CFI = .867, TLI = .836 CFA: bifactor model (one general and three specific factors), with 14 items (less Item 14) RMSEA = .025, CFI = .961, TLI = .944	Model-based reliability of bifactor model, with 14 items (less Item 14) Total: ECV = .43, $\omega_h = .56$; Factor 1: ECV = .18, $\omega_h = .37$, Factor 2: ECV = .19, $\omega_h = .44$, Factor 3: ECV = .19, $\omega_h = .37$, PUC = .71 Simulation with α, with 14 items (less Item 14) Total: α = .73, Factor 1: α = .54, Factor 2: α = .63, Factor 3: α = .66 Correlations between factors: General to Factor 1: $r_s = .68$ General to Factor 2: $r_s = .69$ General to Factor 3: $r_s = .73$ Factor 1 to Factor 2: $r_s = .22$ Factor 1 to Factor 3: $r_s = .33$ Factor 2 to Factor 3: $r_s = .24$	14 items (less item 14; $n = 31$) Total: ICC = .95, Factor 1: ICC = .9, Factor 2: ICC = .88, Factor 3: ICC = .84 (over 1 month)
Dapap and Andu (2021; Nigeria)	Health workers across health professions in a tertiary health setting	308		Total: α = .98, Factor 1: α = .97, Factor 2: α = .95, Factor 3: α = .98	

(*table continues*)

Table 1 (*continued*)

Researcher, year of publication, and country	Research population	*N*	Result		
			Factor structure	Internal consistency	Test–retest reliability
Zuaboni et al. (2021; Switzerland and Germany)	Health workers in general psychiatric inpatient wards	392	EFA: three factors, with 15 items—Items 10, 14, 18, and 19 did not load well unto respective factors CFA: three factors, with 15 items SRMR = .05, RMSEA = .04, CFI = .92	Total: α = .74, Factor 1: α = .62, Factor 2: α = .55, Factor 3: α = .69	
Carrara et al. (2023; Brazil)	Health workers across health professions in family health settings	199	EFA: three factors, with 16 items (fourth factor labeled as "negative views") CFA: four factors, with 20 items SRMR = .067 RMSEA = .043 CFI = .837 TLI = .811 AIC = 10514.842 BIC = 10730.521 SABIC = 10521.446 CFA: four factors, with 16 items SRMR = .06 RMSEA = .043 CFI = .878 TLI = .85 AIC = 8356.547 BIC = 8533.012 SABIC = 8361.951 CFA: five factors, with 16 items SRMR = .057 RMSEA = .041 CFI = .902 TLI = .868 AIC = 8353.067 BIC = 8558.942 SABIC = 8359.371 CFA: one factor, with 16 items SRMR = .076 RMSEA = .069 CFI = .671 TLI = .62 AIC = 8435.081 BIC = 8591.938 SABIC = 8439.884 CFA: five factors, with 16 items SRMR = .06 RMSEA = .041 CFI = .889 TLI = .866 AIC = 8352.578 BIC = 8522.507 SABIC = 8357.781	Model-based reliability of bifactor model, with 14 items (less Item 14) Total: ω = 0.78; Factor 1: ω = 0.68, Factor 2: ω = 0.91, Factor 3: ω = 0.55, Factor 4: ω = 0.60 General OMS factor: ω_h = 0.58, Factor 1: ω_h = 0.25, Factor 2: ω_h = 0.11, Factor 3: ω_h = 0.04, Factor 4: ω_h = 0.2	

(*table continues*)

Table 1 (*continued*)

Researcher, year of publication, and country	Research population	*N*	Result		
			Factor structure	Internal consistency	Test–retest reliability
Őri et al. (2023; 32 European countries)	Health workers and health workers in training, from child and adult psychiatry	4,245	ESEM = bifactor model (one general and three specific factors), with 15 items WRMR = 1.2 RMSEA = .045 CFI = .981 TLI = .96 One factor, with 15 items RMSEA = .078–.172 CFI = .579–.902 TLI = .509–.886 Three factors, with 15 items RMSEA = .054–.124 CFI = .726–.966 TLI = .669—.959	Model-based reliability of bifactor model, with 15 items Total: PUC = .705, ECV = .682, ω_h = .746 Correlations between factors: General to Factor 1: r_s = .79 General to Factor 2: r_s = .68 General to Factor 3: r_s = .79 Factor 1 to Factor 2: r_s = .29 Factor 1 to Factor 3: r_s = .51 Factor 2 to Factor 3: r_s = .31	
Valdivia Ramos et al. (2023; Mexico)	Health workers and health workers in training across different health professions in family medicine settings and academic institutions	556	CFA: three factors, with 15 items SRMR = .054, RMSEA = .05, CFI = .97, TLI = .962, NFI = .95 PNFI = .742	Total: α = .73 and ω = .76, Factor 1: α = .61, Factor 2: α = .51, Factor 3: α = .60	
Cid et al. (2024; Canada)	Health workers in training, from pharmacy	91		Total, with 15 items: Opioid dependency—α = .9 Opioid use disorder—α = .79 Opioid addiction—α = .63 Opioid misuse/use disorder—α = .75	

Note. The data presented in the table are aggregated, rather than data for each group in that study. OMS-HC = Opening Minds Scale for Health Care Providers; PCA = principal component analysis; α = Cronbach's α; ICC = intraclass correlation coefficient; EFA = exploratory factor analysis; CFA = confirmatory factor analysis; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker–Lewis index; *r* = Pearson correlation coefficient; ESEM = exploratory structural equation modeling; SRMR = standardized root-mean-squared residual; WRMR = weighted root-mean-squared residual; PSI = Person separation index; χ^2 = chi square; SEM = structural equation modeling; ECV = explained common variance; PUC = percent of uncontaminated correlations; AIC = Akaike information criteria; BIC = Bayesian information criteria; SABIC = sample size-adjusted BIC; r_s = Spearman's rho; NFI = normalized fit index; PNFI = parsimony normed fit index.

absolute and one relative fit index (Happell et al., 2019; Sapag et al., 2019; Zuaboni et al., 2021).

A further limitation pertains to the predominant focus on internal consistency over test–retest reliability in all but two studies (Destrebecq et al., 2018; Őri et al., 2020). This emphasis may be problematic as internal consistency can be lower for constructs susceptible to change over time (Kruyen, 2012), such as stigmatizing attitudes toward people experiencing mental distress. Moreover, the reliance on Cronbach's *a* for assessing reliability, while common, may be inappropriate for multidimensional measures like the OMS-HC, as it assumes a unidimensional construct where all items contribute equally to the measure (Schmitt, 1996; Sijtsma, 2009).

These inconsistencies in methodological approaches employed to evaluate the OMS-HC are a source of concern, particularly given the often-limited reporting of these methods in the validation studies and their potential lack of robustness. These variabilities may contribute to the conflicting evidence regarding a measure's validity (Hussey & Hughes, 2020). Consequently, when measures with uncertain validity are used to assess interventions, the overall credibility of the evaluation can be compromised (Flake & Fried, 2020). These measurement practices do not support the previous call for thoughtful development and evaluation in stigma research.

Shared Conceptual Theme

In addition to the factor analyses and reliability estimations, a notable finding from the OMS-HC initial evaluation was identifying an underlying conceptual framework that links the study's subscale factors. This interconnected system of variables resembles a nomological net (Cronbach & Meehl, 1955) or reflects a "shared conceptual theme" (Modgill et al., 2014, p. 8). That is, there is a potential for negative beliefs and emotions to translate into harmful behavioral intentions like reduced disclosure and interpersonal distancing. These convergent associations support the view that self-stigma may result from the internalization of public stigma (Corrigan & Watson, 2002), whereas this theme finds backing in two of the subsequent validations of the OMS-HC (Őri et al., 2020, 2023). However, this confirmatory work has not been preregistered in an online protocol. Without this detail, it is difficult to determine whether robust support for the shared conceptual theme is found.

The Present Study

This study aimed to conduct an open and rigorous evaluation of the psychometric properties of the OMS-HC-15 (including the OMS-HC-14) by replicating the work of Őri et al. (2020). The latter represents the most comprehensive examination of the measure to date. The goal was to enhance understanding of the measure's internal structure. Őri et al. (2020) investigated the known structure, unidimensionality, internal consistency, test–retest reliability, and shared conceptual theme of the OMS-HC-15. They also validated the proffered evidence on OMS-HC-14 after obtaining a poor factor loading of Item 14 on all three factors. However, their evaluation was constrained by the absence of a stable fit index in their confirmatory analyses, and their protocols were not shared in advance in an accessible repository.

The present study adopted a partial replication approach to address these methodological limitations and advance open science principles. This research aimed to enhance transparency, reproducibility, and collaboration by openly sharing data, code, and materials. Many metrics, cutoffs, and decision-making processes aligned with best practice recommendations (Hussey & Hughes, 2020). Nevertheless, the study hypotheses were derived from the conclusions of the 2020 evaluation.

Hypotheses

The present study assessed three measurement models: correlated three-factor, unidimensional, and bifactor models. Based on supplemental fit indices, Őri et al. (2020) identified their bifactor model as the best fitting structure, while the three-factor model did not meet the acceptable criteria. Consequently, it was hypothesized that the data would demonstrate a "good" fit with the bifactor model (H1) and "mixed" fit with the three-factor model (H2). By contrast, the data would have a "poor" fit with the unidimensional model (H3).

Despite offering an alternative explanation for the OMS-HC's internal structure, the bifactor model had limitations in model-based reliability. It was then hypothesized that the bifactor model would exhibit a percent of uncontaminated correlations (PUC) $< .8$, an explained common variance $> .6$, and an $\omega_h > .7$ (H4). Additionally, it was anticipated that the OMS-HC-14 and OMS-HC-15 scores would have a 95% confidence interval lower bound for internal consistency of $\omega_t \geq .7$ (H5). At the same time, the subscales would show internal consistency of $\omega_t < .7$ (H6).

Based on Őri et al.'s (2020) study reporting good to excellent test–retest reliability for both general and specific factors, it was hypothesized that the 95% confidence interval lower bound for test–retest reliability of the OMS-HC-14 and OMS-HC-15 total and subscale scores would be $r \geq .7$ (H7). Additionally, given the strong correlations between the general and specific factors and weaker correlations among the specific factors for the "shared conceptual theme" in the 2020 study, it was hypothesized that correlations between the general and specific factors would exceed $r_s > .59$ (H8), while correlations among the specific factors would be $r_s \leq .39$ (H9).

Method

This article focuses on the findings related to the stigmatizing attitudes measure within a more extensive study investigating measures of recovery and stigma. The study employed an open science approach by preregistering the detailed protocol on the Open Science Framework (Foster & Deardorff, 2017; see https://osf.io/tuh49/). Data collection was undertaken from January to February 2022.

Design

Data were collected using a cross-sectional online Qualtrics survey, with a subsample of participants completing a follow-up survey 2 weeks later. To ensure adequate power (90%) to detect misspecification of the OMS-HC-15's three-factor structure, an a priori power analysis based on a root-mean-square error of approximation of .045—similar to that obtained in the correlated three-factor model by Őri et al. (2023)—determined a minimum sample size of 286 participants (includes 10% to account for potential exclusions). A separate analysis indicated that a minimum of 19 participants (including an extra 10%) were required for the test–retest reliability analysis with 90% power.

Participants

Participants were recruited via Prolific, a diverse online platform offering paid surveys (Prolific, 2025). While the participant pool is global, it predominantly comprises individuals from the United States and the United Kingdom. Prolific verifies participants through a four-step process—email, phone, ID verification, and a trial study—to confirm their identity and ensure they are who they claim to be (Prolific, 2022). To ensure further data quality, participants were required to have a high Prolific approval rating (≥95%), be fluent in English, be aged 18 or older, and be employed in the medical sector (though not necessarily in mental health). After excluding 34 participants outside the medical field, 286 completed the initial survey. A subsample of 19 participants completed a follow-up survey 2 weeks later to assess test–retest reliability. A 2-week interval is commonly used in test–retest reliability studies to balance the risk of recall bias with potential changes in the measured construct (Polit, 2014).

Measures

Participants completed an online Qualtrics survey. The survey collected demographic data (age, gender, and ethnicity) for descriptive purposes. Subsequently, participants completed the OMS-HC-15, a self-report measure assessing attitudes and behavioral intentions toward individuals with lived experience of mental distress (Modgill et al., 2014). This 15-item scale employs a 5-point Likert format (1 = *strongly disagree*, 5 = *strongly agree*). Total and subscale scores were derived by summing item responses, with reverse scoring applied to Items 3, 8, 9, 10, 11, 15, and 19. Total scores range from 15 to 75, with Factor 1 (Items 1, 12, 13, 14, 18, and 20) ranging from 6 to 30, Factor 2 (Items 4, 6, 7, and 10) from 4 to 20, and Factor 3 (Items 3, 8, 9, 17, and 19) from 5 to 25. Lower scores indicate lesser stigmatizing attitudes and behavioral intentions. Scores near 15 indicate lesser stigmatizing attitudes and behavioral intentions, while scores approaching 75 reflect greater stigmatizing attitudes and behavioral intentions. An attention check question was included toward the survey's conclusion to enhance data quality. The broader survey encompassed an additional measure, totaling 22 items.

Procedure

Participant recruitment occurred via Prolific Academic. Potential participants accessed a Qualtrics survey link to obtain further study information before providing consent and completing the survey. Participants received a £1.25 payment for their time, which was considered insufficient to constitute an inducement. A randomly selected subset of participants received payment to complete a follow-up survey after 2 weeks. Qualtrics' "Prevent multiple submissions" feature was enabled to avoid duplicate submissions during the initial data collection but was disabled for follow-up. Inclusion criteria encompassed providing informed consent, being 18 or older, and correctly answering an attention check question. According to a New Zealand university's research ethics guideline, the study was deemed low risk and exempt from formal ethics committee review (Ethics Notification Number: 4000025270). All research procedures adhered to the university's research code of ethics.

Data Analyses

Seven participants had one to two missing items within the broader survey, while one missed an item during the retest. Imputation was employed to address these missing data points. The complete, untransformed data set exhibited multivariate nonnormality, consistent with previous research analyzing OMS-HC-15 scores nonparametrically (e.g., Őri et al., 2020). However, to avoid potential spurious findings associated with mixed validity support (Flake & Fried, 2020), the planned statistical analyses were maintained despite whatever data distribution was ascertained. De-identified data were deposited on the Open Science Framework.

Given the established factor structure of the OMS-HC-15, confirmatory factor analyses were conducted using a combination of R (R Core Team, 2021) and RStudio (Posit Team, 2023). Unidimensional, bifactor, and three-factor models were specified and fitted to the data. To accommodate the nonnormal distribution, maximum likelihood estimation with robust standard errors and mean- and variance-adjusted adjustments was utilized (Rhemtulla et al., 2012).

The study employed established criteria to assess model fit from Hu and Bentler (1999): standardized root-mean-square residual ≤ .09, root-mean-square error of approximation ≤ .06, Tucker Lewis index ≥ .95, and comparative fit index ≥ .95. Interpretation of model fit against the above criteria was based on convention (see Hussey & Hughes, 2020). Therefore, Hypothesis 1 regarding support for a specific model would be confirmed if it satisfied all four criteria. Hypothesis 2, which compared the bifactor model with the three-factor model, would be upheld if the bifactor model met more criteria than the three-factor model. Conversely, Hypothesis 3 would be supported if none of the criteria were met.

Reliability evidence was assessed using McDonald's omega (ω_t) and Pearson's correlation coefficient (r). Coefficients exceeding or equaling .7 were considered indicative of good reliability (Nunnally, 1978). McDonald's ω is preferred for shorter measures (Ziegler et al., 2014), as it accommodates potential variations in item structure while acknowledging a shared underlying construct (congeneric). In addition, when the percentage of uncontaminated correlation values exceeds 0.80, the general explained common variance values become less critical for predicting bias (Őri et al., 2020). If the percentage of uncontaminated correlations values is below 0.80, an explained common variance above 0.60 and ω_h above 0.70 suggest that multidimensionality is not severe enough to rule out treating the instrument as mainly unidimensional (Reise et al., 2013).

Results

The study sample consisted primarily of young adults (75% aged 18–30), with a majority identifying as female (71%) and European (63%). More than half (53%) were residents of the United Kingdom. Further details on participant demographics (age, ethnicity, gender) can be found in the Supplemental Materials. Conversely, Table 2 indicates that the OMS-HC-15 scores generally reflected lesser stigmatizing attitudes.

The evaluation explored three models for the OMS-HC-15: three-factor, unidimensional, and bifactor. The three-factor model exhibited an acceptable fit based on only two fit indices (see Table 3). The unidimensional model met the threshold for one fit index (standardized root-mean-square residual). Attempts to fit the bifactor model resulted in convergence failure, whereas rerunning the

Table 2
Descriptive Statistics of the Total and Subscale OMS-HC Scores

Scale	*M*	*SD*	*Mdn*	Min	Max	Possible range
OMS-HC-15 total	31.86	7.80	32	15	62	15–75
Factor 1	12.79	3.73	13	6	25	6–30
Factor 2	10.31	3.16	10	4	18	5–25
Factor 3	8.76	2.83	9	5	20	4–20
OMS-HC-14 total	29.67	7.26	30	14	57	14–70
Factor 1 (minus Item 14)	10.60	3.17	10	5	20	5–25

Note. Factor 1 = Attitude; Factor 2 = Disclosure and Help-seeking; Factor 3 = Social Distance; OMS-HC = Opening Minds Scale for Health Care Providers; Min = minimum; Max = maximum.

analysis with the Broyden–Fletcher–Goldfarb–Shanno algorithm yielded negative latent variances for Factor 1 (Attitude). As such, standardized coefficients for model-based reliability calculations were unavailable. Hypotheses 1–4 were then rejected.

Lower bounds of the 95% confidence intervals for total and subscale scores (except attitude without Item 14) regarding internal consistency exceeded 0.7 (see Table 4), supporting Hypothesis 5. However, Hypothesis 6 was not supported. Test–retest reliability estimates were negative, contradicting Hypothesis 7.

Correlations between the general and specific factors in the OMS-HC subscale scores (excluding Item 14) surpassed correlations among specific factors (see Table 5). As a result, Hypothesis 8 was supported, whereas Hypothesis 9 was rejected.

Discussion

This study is the first evaluation of the OMS-HC's factorial structure and reliability that employed preregistration and sharing of Supplemental Materials, which supports transparency and replicability in psychological measurement. It incorporated McDonald's ω coefficient and transparent model fit criteria and prioritized test–retest reliability. This study replicated part of the findings of Őri et al. (2020), in which three models were tested, retest reliability and internal consistency were estimated, and factor relationships were assessed.

Confirmatory Factor Analyses

Despite mixed results, the correlated three-factor structure best fits the present study data. These findings align more with those from the Chilean (Sapag et al., 2019), rather than the Hungarian, study where only root-mean-square error of approximation met the criteria (Őri et al., 2020). A potential explanation could be that the measure might capture an additional stigma dimension (i.e., Anticipated Stigma or Negative Views; see Carrara et al., 2023; Fox et al., 2018) besides the known three factors proposed by its developers. On the other hand, Item 1 in the same model could be reevaluated to improve fit (see path diagram in the Supplemental Materials), as observed in the Singaporean study (Chang et al., 2017).

The unidimensional model exhibited poor fit, aligning with previous research (Carrara et al., 2023; Happell et al., 2019; Őri et al., 2020, 2023). This finding signals that the total OMS-HC score might not be a singular, robust stigma indicator. Simply put, the measure captures multiple dimensions of stigma, and combining them into a single number could obscure the underlying complexities of the central construct.

The bifactor model proposed by the European validation studies (Őri et al., 2020, 2023) did not yield fit indices nor model-based reliability estimates in the current evaluation. This outcome raises concerns about its applicability for representing overall stigmatizing attitudes or individual subscales. The present study's samples may explain the contrasting results, which are from the broader health sector and mostly from the United Kingdom. Earlier studies providing evidence of the bifactor model used samples from a single discipline (i.e., psychiatry or medicine) and different European countries. Furthermore, as this study's a priori sample size calculation was designed to detect model misspecification, there was no ability to account for the requirements of measurement invariance testing to detect differences in model measurement across groups defined by place of origin. As a result, differences in model fit between this study and the European studies may be due to random sample variations. Sampling variabilities can be associated with improper solutions in confirmatory factor analysis (Bollen, 1987).

The lack of good fit between the measurement models and the current data sets, primarily from young survey participants in the United Kingdom (and then the United States), may have resulted from items in the measures used to develop the OMS-HC (see Kassam et al., 2012) that no longer reflect evolving understandings of stigma. For instance, the "Opinions about Mental Illness in the Personnel of Two Large Mental Hospitals" (Cohen & Struening, 1962) included deinstitutionalization items that do not apply nowadays. Involuntary admissions in acute care are prevalent (Daguman et al., 2024). However, many mental health services worldwide, including clinical psychology services, have become organized to work in a recovery-oriented fashion (Taylor & Gordon, 2022). The Recovery Attitudes Questionnaire (Borkin et al., 2000)

Table 3
Confirmatory Factor Analyses on the Opening Minds Scale for Health Care Providers

Model	χ^{2a}	χ^2/df	*df*	CFI	TLI	RMSEA [95% CI]	SRMR
Bifactor (14 items)							
Three factors (15 items)	77.950	1.84	42	.754	.897	.054 [.042, .066]	.057
One-factor (15 items)	126.104	3.01	42	.418	.755	.084 [.073, .095]	.080

Note. χ^2 = chi-square statistic; χ^2/df = chi-square to degrees of freedom ratio; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis fit index; RMSEA = root-mean-square error of approximation; CI = confidence interval; SRMR = standardized root-mean-squared residual.
[a] For all χ^2 tests, $\rho \leq .001$.

Table 4
Reliability Estimation on the OMS-HC

Scale	Number of items	α	ω_t	ω_h	r^a
OMS-HC-15	15	.84 [.73, .92]	.86 [.86, .98]	.72 [.20, .74]	−.24 [−.63, .24]
Factor 1	6	.73 [.41, .85]	.77 [.76, .97]	.60 [.33, .77]	−.20 [−.60, .28]
Factor 2	5	.70 [.36, .86]	.76 [.75, .94]	.62 [.18, .89]	−.46 [−.75, —]
Factor 3	4	.74 [.64, .91]	.77 [.71, .97]	.67 [.57, .83]	.21 [−.27, .61]
OMS-HC-14	14	.82 [.75, .90]	.85 [.84, .97]	.45 [.23, .79]	−.29 [−.66, .19]
Factor 1	5	.69 [.01, .75]	.73 [.61, .90]	.63 [.25, .78]	−.29 [−.66, .19]

Note. α = Cronbach's α; ω_t = McDonald's omega total; ω_h = McDonald's omega hierarchical; r = Pearson's correlation coefficient; Factor 1 = Attitude; Factor 2 = Disclosure and Help-seeking; Factor 3 = Social Distance. Values presented in square brackets represent the lower and upper bounds of the 95% confidence intervals, respectively; OMS-HC = Opening Minds Scale for Health Care Providers.
[a] Test–retest interval was 2 weeks.

has also faced criticism for its cultural inappropriateness outside the United States and mixed evidence of validity (Daguman & Taylor, 2024). In the United Kingdom, cultural and generational differences can influence help seeking, with age often perceived as a barrier to a health worker's self-disclosure (Zaman et al., 2022). In the United States, persistent stigma toward a range of mental distress diagnoses and rising fears of violence in schizophrenia suggest a rethink on stigma reduction strategies, although age and generational shifts have improved public stigma around major depression (Pescosolido et al., 2021). These changes support the need to regularly review and update measures of attitudes toward stigma to keep them relevant and accurate in changing social and professional contexts—an intellectual and practical challenge in measuring stigma and making comparisons over time.

Reliability Estimation

The OMS-HC-15's total and subscale scores demonstrated strong internal consistency, with estimates and 95% confidence interval lower bounds exceeding the set cutoff. These results align with the Italian study (Destrebecq et al., 2018) and surpass the developers' findings. Nevertheless, with the factorial validity findings of the present study, the use of highly reliable total scores may need reconsideration. Even if reliability is often (wrongly) seen as the only proof of a valid measure by many in the literature, inferences from scores require broader validity evidence, as both reliability and validity reflect the score's quality, not the instrument itself (Cook & Beckman, 2006).

Unlike the sole other study examining its stability over time (Destrebecq et al., 2018), the OMS-HC-15 retest reliability shows concerning negative values for both total and subscale scores. However, it is argued that these are unlikely due to the 2-week retest interval, which is standard in health research (Polit, 2014). Additionally, these negative coefficients likely do not reflect deliberate intervening effort or a practice effect, as such factors would typically lead to improved scores on the retest—not the decrease observed here. Furthermore, the planned sample size makes limitations in that area improbable. Nevertheless, the findings do not make the case for testing with larger sample size to be disregarded, especially where it is economically possible.

Table 5
Correlation Coefficients Between the Opening Minds Scale for Health Care Providers Total and Subscale Scores

Relationship	r_s
Total score and Factor 1	.83 [.79, .87]
Total score and Factor 2	.75 [.69, .80]
Total score and Factor 3	.76 [.71, .81]
Factors 1 and 2	.45 [.34, .54]
Factors 1 and 3	.55 [.46, .62]
Factors 2 and 3	.32 [.22, .42]

Note. Numbers presented in square brackets are the lower and upper bounds of the 95% confidence interval. r_s = Spearman's correlation coefficient; Factor 1 = Attitude; Factor 2 = Disclosure and Help-seeking; Factor 3 = Social Distance.

Shared Conceptual Theme

The support in the present study for the notion that stereotype, prejudice, and discrimination are interrelated is statistically significant yet weak to moderate, consistent with findings from previous validations (Modgill et al., 2014; Őri et al., 2020, 2023). A note of caution is then due here because the present study also did not confirm that the OMS-HC scores accurately reflect these three stigma mechanisms nor is the total score sufficient to represent the overall stigma construct. Tentativeness in communicating this finding aligns with the view that validating individual relationships within a theory requires first validating the overall theory (Hagger et al., 2017). In other words, even though the statistically significant relationships may represent the interrelatedness of stigma mechanisms, the *in*validity of the total score—supposedly representing the overarching stigma construct in diverse contexts—suggests that the concept of these three mechanisms may not accurately reflect their intended theoretical roles. Therefore, the current evidence on the shared conceptual theme does not strongly support the notion that self-stigma necessarily results from internalizing public stigma, although it remains a possibility.

Implications for Practice

The psychometric profile of the OMS-HC found in the present study presents a unique characteristic: high internal consistency alongside low test–retest reliability. Drawing on the established distinction between trait and state constructs (Widaman et al., 2011),

this profile suggests that the measure might capture fluctuating attitudes toward mental distress stigma rather than enduring traits. In measurement theory, instruments used for evaluation ideally demonstrate responsiveness to change, disparate from those for discrimination that benefit from high test–retest reliability (Mokkink et al., 2010). Although the original development and evaluation of the OMS-HC was silent on whether it measures attitudes as a state or trait, its recently found profile makes it a strong candidate for assessing short-term attitude change, as in the immediate effectiveness of antistigma interventions. Conversely, its low test–retest reliability questions its use in tracking sustained attitude change. Further research is crucial to determine if the measure can be refined or if a complementary measure targeting stable traits is necessary.

Implications for Research

The present study provides value by utilizing online samples from the medical sector, which broadens the understanding of stigma within its broader social context (Thornicroft et al., 2022). However, the study lacked details regarding the specific occupations within the online "medicine" sector sample (e.g., doctors, nurses, administrative staff). This lack of granularity may hinder the ability to stratify attitudes based on potential variations in stigma perceptions across different health worker roles through measurement invariance testing, which can also help robustly identify differences in measurement models across gender-defined groups. This limitation also does not support the recognition of potential stigma by association against a specific profession, where negative attitudes are directed at a profession due to its connection with other people's characteristics (Njaka et al., 2023). Understanding stigma by association helps identify whether stigma linked to a specific profession, rather than public stigma, contributes to workers' self-stigma and informs targeted interventions.

What also remains unclear is whether lesser stigmatizing attitudes can be found at all levels of a health service. Despite offering the OMS-HC to a vast pool, most of the gathered measures are from workers at the early stage of their professional experience. The generalizability of the findings is then limited. It may be helpful to communicate that countering stigma in the everyday practice of care provision is an organizational effort (Knaak et al., 2017). If services are to support service users inclusively, lesser stigmatizing attitudes will then need to be seen across the service, including workers in senior positions.

Finally, the recruitment platform employed in the study serves participants from the United Kingdom and the United States and the current sample composition reflects this bias, with a majority of European health workers. Consequently, the findings offer a minimal contribution to the discourse on health research within non-Western contexts. This limitation underscores the importance of culturally adapting stigma measurement tools, as emphasized by previous studies on the OMS-HC (Carrara et al., 2023; Sapag et al., 2019; Valdivia Ramos et al., 2023). Cultural variations can significantly influence perceptions of stigma, necessitating thoughtful validation of measures across diverse cultural settings (Yang et al., 2013).

Conclusions

This study assessed the OMS-HC's factorial structure and reliability. The mixed results of confirmatory factor analyses suggest that the measure may capture additional dimensions of stigma beyond the proposed factors. The reliability estimation findings imply that, while the OMS-HC can reflect short-term attitude changes, its low stability questions its suitability for tracking long-term stigma changes. Future research needs to consider addressing these limitations and exploring the measure's applicability across diverse populations.

References

Bollen, K. A. (1987). Outliers and improper solutions. *Sociological Methods & Research*, *15*(4), 375–384. https://doi.org/10.1177/0049124187015004002

Borkin, J. R., Steffen, J. J., Ensfield, L. B., Krzton, K., Wishnick, H., Wilder, K., & Yangarber, N. (2000). Recovery Attitudes Questionnaire: Development and evaluation. *Psychiatric Rehabilitation Journal*, *24*(2), 95–102. https://doi.org/10.1037/h0095112

British Psychological Society. (2015). *Guidelines on language in relation to functional psychiatric diagnosis*. https://criticalmhnursing.org/wp-content/uploads/2015/09/guidelines_on_language_web.pdf

Carrara, B. S., Sanches, M., Bobbili, S. J., de Godoy Costa, S., de Sousa, Á. F. L., de Souza, J., & Ventura, C. A. A. (2023). Validation of the Opening Minds Scale for Health Care Providers (OMS-HC): Factor structure and psychometric properties of the Brazilian version. *Healthcare*, *11*(7), Article 1049. https://doi.org/10.3390/healthcare11071049

Chang, S., Ong, H. L., Seow, E., Chua, B. Y., Abdin, E., Samari, E., Teh, W. L., Chong, S. A., & Subramaniam, M. (2017). Stigma towards mental illness among medical and nursing students in Singapore: A cross-sectional study. *BMJ Open*, *7*(12), Article e018099. https://doi.org/10.1136/bmjopen-2017-018099

Cid, A., Patten, A., Beazely, M. A., & Grindrod, K. (2024). Adapting the Opening Minds Stigma Scale for Healthcare Providers to measure opioid-related stigma. *Pharmacy*, *12*(4), Article 105. https://doi.org/10.3390/pharmacy12040105

Cohen, J., & Struening, E. L. (1962). Opinions about mental illness in the personnel of two large mental hospitals. *The Journal of Abnormal and Social Psychology*, *64*(5), 349–360. https://doi.org/10.1037/h0045526

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine*, *119*(2), 166.e7–166.e16. https://doi.org/10.1016/j.amjmed.2005.10.036

Corrigan, P. W. (2005). *On the stigma of mental illness: Practical strategies for research and social change*. American Psychological Association. https://doi.org/10.1037/10887-000

Corrigan, P. W., & Watson, A. C. (2002). The paradox of self-stigma and mental illness. *Clinical Psychology: Science and Practice*, *9*(1), 35–53. https://doi.org/10.1093/clipsy.9.1.35

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Daguman, E., IV, & Taylor, J. E. (2024). Do health workers believe that personal recovery is possible? An evaluation of the Recovery Attitudes Questionnaire. *Journal of Psychosocial Rehabilitation and Mental Health*. Advance online publication. https://doi.org/10.1007/s40737-024-00442-y

Daguman, E., IV, Hutchinson, M., & Lakeman, R. (2024). Uncovering complexities in reducing aggression, conflict and restrictive practices in acute mental healthcare settings: An overview of reviews. *International Journal of Mental Health Nursing*, *33*(6), 1666–1686. https://doi.org/10.1111/inm.13376

Dapap, D. D., & Andu, M. D. (2021). Validation of opening minds scale for health care providers (OMS-HC). *International Journal of Medical and Health Research*, *7*(12), 26–30.

Destrebecq, A., Ferrara, P., Frattini, L., Pittella, F., Rossano, G., Striano, G., Terzoni, S., & Gambini, O. (2018). The Italian version of the Opening Minds Stigma Scale for Healthcare Providers: Validation and study on a

sample of bachelor students. *Community Mental Health Journal*, *54*(1), 66–72. https://doi.org/10.1007/s10597-017-0149-0

Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, *3*(4), 456–465. https://doi.org/10.1177/2515245920952393

Foster, E. D., & Deardorff, A. (2017). Open science framework (OSF). *Journal of the Medical Library Association*, *105*(2), Article 203. https://doi.org/10.5195/jmla.2017.88

Fox, A. B., Earnshaw, V. A., Taverna, E. C., & Vogt, D. (2018). Conceptualizing and measuring mental illness stigma: The mental illness stigma framework and critical review of measures. *Stigma and Health*, *3*(4), 348–376. https://doi.org/10.1037/sah0000104

Gupta, S., Kumar, A., Kathiresan, P., Pakhre, A., Pal, A., & Singh, V. (2024). Mental health stigma and its relationship with mental health professionals—A narrative review and practice implications. *Indian Journal of Psychiatry*, *66*(4), 336–346. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_412_23

Hagger, M. S., Gucciardi, D. F., & Chatzisarantis, N. L. D. (2017). on nomological validity and auxiliary assumptions: The importance of simultaneously testing effects in social cognitive theories applied to health behavior and some guidelines. *Frontiers in Psychology*, *8*, Article 1933. https://doi.org/10.3389/fpsyg.2017.01933

Happell, B., Platania-Phung, C., Scholz, B., Bocking, J., Horgan, A., Manning, F., Doody, R., Hals, E., Granerud, A., Jan van der Vaart, K., Allon, J., Lahti, M., Pulli, J., Vatula, A., Ellilä, H., Griffin, M., Russell, S., MacGabhann, L., Bjornsson, E., … Biering, P. (2019). Assessment of the Opening Minds Scale for use with nursing students. *Perspectives in Psychiatric Care*, *55*(4), 661–666. https://doi.org/10.1111/ppc.12393

Heim, E., Kohrt, B. A., Koschorke, M., Milenova, M., & Thornicroft, G. (2020). Reducing mental health-related stigma in primary health care settings in low- and middle-income countries: A systematic review. *Epidemiology and Psychiatric Sciences*, *29*, Article e3. https://doi.org/10.1017/S2045796018000458

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hussey, I., & Hughes, S. (2020). Hidden invalidity among 15 commonly used measures in social and personality psychology. *Advances in Methods and Practices in Psychological Science*, *3*(2), 166–184. https://doi.org/10.1177/2515245919882903

Kassam, A., Papish, A., Modgill, G., & Patten, S. (2012). The development and psychometric properties of a new scale to measure mental illness related stigma by health care providers: The Opening Minds Scale for Health Care Providers (OMS-HC). *BMC Psychiatry*, *12*(1), Article 62. https://doi.org/10.1186/1471-244X-12-62

Knaak, S., Mantler, E., & Szeto, A. (2017). Mental illness-related stigma in healthcare: Barriers to access and care and evidence-based solutions. *Healthcare Management Forum*, *30*(2), 111–116. https://doi.org/10.1177/0840470416679413

Kruyen, P. M. (2012). *Using short tests and questionnaires for making decisions about individuals: When is short too short?* Ridderprint.

Link, B. G., & Phelan, J. C. (2001). Conceptualizing Stigma. *Annual Review of Sociology*, *27*(1), 363–385. https://doi.org/10.1146/annurev.soc.27.1.363

Livingston, J. D. (2020). *Structural stigma in health-care contexts for people with mental health and substance use issues*. Mental Health Commission of Canada.

Modgill, G., Patten, S. B., Knaak, S., Kassam, A., & Szeto, A. C. H. (2014). Opening Minds Stigma Scale for Health Care Providers (OMS-HC): Examination of psychometric properties and responsiveness. *BMC Psychiatry*, *14*(1), Article 120. https://doi.org/10.1186/1471-244X-14-120

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Na, J. J., Park, J. L., Lkhagva, T., & Mikami, A. Y. (2022). The efficacy of interventions on cognitive, behavioral, and affective public stigma around mental illness: A systematic meta-analytic review. *Stigma and Health*, *7*(2), 127–141. https://doi.org/10.1037/sah0000372

Njaka, S., Ogbodo, B. A., Okpua, N. C., & Uwa, A. N. (2023). A systematic review on prevalence and perceived impacts of associative stigma on mental health professionals. *International Journal of Africa Nursing Sciences*, *18*, Article 100533. https://doi.org/10.1016/j.ijans.2023.100533

Nunnally, J. C. (1978). An overview of psychological measurement. In B. B. Wolman (Ed.), *Clinical diagnosis of mental disorders: A handbook* (pp. 97–146). Springer.

O'Hagan, M. (2018). *Wellbeing manifesto for Aotearoa New Zealand: A submission to the government inquiry into mental health and addiction*. https://www.wellbeingmanifesto.nz/

Őri, D., Rózsa, S., Szocsics, P., Simon, L., Purebl, G., & Győrffy, Z. (2020). Factor structure of The Opening Minds Stigma Scale for Health Care Providers and psychometric properties of its Hungarian version. *BMC Psychiatry*, *20*(1), Article 504. https://doi.org/10.1186/s12888-020-02902-8

Őri, D., Szocsics, P., Molnár, T., Bankovska Motlova, L., Kazakova, O., Mörkl, S., Wallies, M., Abdulhakim, M., Boivin, S., Bruna, K., Cabacos, C., Carbone, E. A., Dashi, E., Grech, G., Greguras, S., Ivanovic, I., Guevara, K., Kakar, S., Kotsis, K., … Rózsa, S. (2023). Psychometric properties of the Opening Minds Stigma Scale for Health Care Providers in 32 European countries—A bifactor ESEM representation. *Frontiers in Public Health*, *11*, Article 1168929. https://doi.org/10.3389/fpubh.2023.1168929

Pescosolido, B. A., Halpern-Manners, A., Luo, L., & Perry, B. (2021). Trends in public stigma of mental illness in the US, 1996–2018. *JAMA Network Open*, *4*(12), Article e2140202. https://doi.org/10.1001/jamanetworkopen.2021.40202

Polit, D. F. (2014). Getting serious about test–retest reliability: A critique of retest research and some recommendations. *Quality of Life Research*, *23*(6), 1713–1720. https://doi.org/10.1007/s11136-014-0632-9

Prolific. (2022). *Prolific's data quality just got even better with Onfido ID verification*. Prolific. https://www.prolific.com/resources/onfido-id-verification?utm_source=chatgpt.com

Prolific. (2025). *Academic researchers*. Prolific. https://www.prolific.com/academic-researchers

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, *73*(1), 5–26. https://doi.org/10.1177/0013164412449831

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–373. https://doi.org/10.1037/a0029315

Ruesch, N. (2022). *The stigma of mental illness-E-book: Strategies against discrimination and social exclusion*. Elsevier Health Sciences.

Sapag, J. C., Klabunde, R., Villarroel, L., Velasco, P. R., Álvarez, C., Parra, C., Bobbili, S. J., Mascayano, F., Bustamante, I., Alvarado, R., & Corrigan, P. (2019). Validation of the Opening Minds Scale and patterns of stigma in Chilean primary health care. *PLOS ONE*, *14*(9), Article e0221825. https://doi.org/10.1371/journal.pone.0221825

Sastre-Rus, M., García-Lorenzo, A., Lluch-Canut, M.-T., Tomás-Sábado, J., & Zabaleta-Del-Olmo, E. (2019). Instruments to assess mental health-related stigma among health professionals and students in health sciences: A systematic psychometric review. *Journal of Advanced Nursing*, *75*(9), 1838–1853. https://doi.org/10.1111/jan.13960

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*(4), 350–353. https://doi.org/10.1037/1040-3590.8.4.350

Sijtsma, K. (2009). Reliability beyond theory and into practice. *Psychometrika*, *74*(1), 169–173. https://doi.org/10.1007/s11336-008-9103-y

Taylor, J. E., & Gordon, S. E. (2022). Evaluating service user-led teaching of mental health recovery concepts in clinical psychology training. *Training and Education in Professional Psychology*, *16*(4), 385–393. https://doi.org/10.1037/tep0000377

Posit Team. (2023). *RStudio: Integrated development environment for R*. RStudio PBC.

Thornicroft, G. (2006). *Shunned: Discrimination against people with mental illness*. Oxford University Press.

Thornicroft, G., Sunkel, C., Alikhon Aliev, A., Baker, S., Brohan, E., El Chammay, R., Davies, K., Demissie, M., Duncan, J., Fekadu, W., Gronholm, P. C., Guerrero, Z., Gurung, D., Habtamu, K., Hanlon, C., Heim, E., Henderson, C., Hijazi, Z., Hoffman, C., … Winkler, P. (2022). The Lancet Commission on ending stigma and discrimination in mental health. *The Lancet*, *400*(10361), 1438–1480. https://doi.org/10.1016/S0140-6736(22)01470-2

Valdivia Ramos, H. N., Mora-Rios, J., Natera, G., & Mondragón, L. (2023). Psychometric properties of the Mexican version of the opening minds stigma scale for health care providers (OMS-HC). *PeerJ*, *11*, Article e16375. https://doi.org/10.7717/peerj.16375

van der Maas, M., Stuart, H., Patten, S. B., Lentinello, E. K., Bobbili, S. J., Mann, R. E., Hamilton, H. A., Sapag, J. C., Corrigan, P., & Khenti, A. (2018). Examining the application of the opening minds survey in the community health centre setting. *Canadian Journal of Psychiatry*, *63*(1), 30–36. https://doi.org/10.1177/0706743717719079

Watts, G. (2022). Graham Thornicroft: The struggle against stigma in mental health. *The Lancet*, *400*(10361), Article 1397. https://doi.org/10.1016/S0140-6736(22)01934-1

Widaman, K. F., Little, T. D., Preacher, K. J., & Sawalani, G. M. (2011). On creating and using short forms of scales in secondary research. In K. H. Trzesniewski, M. B. Donnellan, & R. E. Lucas (Eds.), *Secondary data analysis: An introduction for psychologists* (pp. 39–61). American Psychological Association. https://doi.org/10.1037/12350-003

Yang, L. H., Valencia, E., Alvarado, R., Link, B., Huynh, N., Nguyen, K., Morita, K., Saavedra, M., Wong, C., Galea, S., & Susser, E. (2013). A theoretical and empirical framework for constructing culture-specific stigma instruments for Chile. *Cadernos Saúde Coletiva*, *21*(1), 71–79. https://doi.org/10.1590/S1414-462X2013000100011

Zaman, N., Mujahid, K., Ahmed, F., Mahmud, S., Naeem, H., Riaz, U., Ullah, U., & Cox, B. (2022). What are the barriers and facilitators to seeking help for mental health in NHS doctors: A systematic review and qualitative study. *BMC Psychiatry*, *22*(1), Article 595. https://doi.org/10.1186/s12888-022-04202-9

Ziegler, M., Kemper, C. J., & Kruyen, P. (2014). Short scales—Five misunderstandings and ways to overcome them. *Journal of Individual Differences*, *35*(4), 185–189. https://doi.org/10.1027/1614-0001/a000148

Zuaboni, G., Elmer, T., Rabenschlag, F., Heumann, K., Jaeger, S., Kozel, B., Mahlke, C. I., Theodoridou, A., Jaeger, M., & Rüsch, N. (2021). Psychometric evaluation of the German version of the Opening Minds Stigma Scale for Health Care Providers (OMS-HC). *BMC Psychology*, *9*(1), Article 86. https://doi.org/10.1186/s40359-021-00592-9

Received July 24, 2024
Revision received January 29, 2025
Accepted January 30, 2025 ■