



Review article

A systematic review of studies using translated versions of the Attribution Questionnaire (AQ-27) to measure public stigma towards people with schizophrenia

Claire Thirkettle^a, Sheri Oduola^b, Sharon Black^c, Lucy McEntegart^d, Peter Beazley^{a,*}

^a Department of Clinical Psychology and Psychological Therapies, Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK

^b School of Health Sciences, University of East Anglia, UK

^c School of Media, Language and Communication Studies, University of East Anglia, UK

^d Norfolk and Suffolk NHS Foundation Trust, UK



ARTICLE INFO

Keywords:

AQ-27
Translation
Cross-cultural
Stigma
Schizophrenia

ABSTRACT

Background and Objectives: The Attribution Questionnaire (AQ-27) is a widely used measure of public mental illness stigma. The AQ-27 was originally developed in the USA in the English language. Since its inception in 2003, several translations of the measure have been produced. This is the first review to explore the use of translated versions of the AQ-27 to measure stigma towards people with schizophrenia.

Methods: A systematic review was conducted. MEDLINE, PsycInfo and Web of Science were systematically searched between 2003 and 2024. The COSMIN Study Design Checklist was adapted to appraise the quality of the translation processes. Data were extracted relating to measurement properties (reliability and validity) of the translated measures.

Results: Forty-one studies were identified, spanning fifteen countries and eleven languages. Most studies ($n = 26$, 63.4 %) were located in Europe. Twelve original translations of the AQ-27 were identified, of which, four studies were primarily focused on translation and validation of the measure. The Turkish, Italian and Arabic translations were rated highest for methodological quality of the translation process.

Conclusions: Researchers should consider the quality of the methodology used to develop existing translated versions of the AQ-27 before adopting them, as this may have implications for the validity and equivalence of the measure within the target culture. Translation frameworks are available to support the high-quality translation and cross-cultural adaptation of self-report measures.

Introduction

Defining stigma

Across countries and cultures, the psychiatric diagnosis of schizophrenia is associated with a high level of public stigma and experienced discrimination.^{1,2} Stigma has been defined in a variety of ways. Erving Goffman's conceptualisation of stigma as being an 'attribute that is deeply discrediting'³ has been built on by authors such as Link and Phelan⁴ who conceptualise stigma as consisting of several interacting components: the labelling of difference, stereotyping, separation of 'us' and 'them', status loss, and discrimination. Power differences (social, economic and political) are considered crucial to enabling stigmatisation. Other perspectives emphasise the role of culture and the social

context in defining stigma, whereby stigma is thought to pose a threat to one's moral standing within the local social world.^{5,6}

The reduction of stigma, discrimination and human rights violations towards people with mental health difficulties has been identified as a key priority within the WHO Comprehensive Mental Health Action Plan (2013–2030).⁷ Further to this, a recent report by the Lancet Commission outlines eight key recommendations for action worldwide.⁸ Regarding global mental health and stigma reduction, research suggests that cross-cultural variation exists in public stigma.⁹ However, there is limited research taking place outside of the Global North to indicate effective, culturally appropriate strategies for stigma reduction.¹⁰ Research is needed across different countries and cultural settings, including developing countries and the Global South to explore the efficacy and feasibility of methods to address stigma.^{2,11} Additionally, a

* Corresponding author.

E-mail address: p.beazley@uea.ac.uk (P. Beazley).

<https://doi.org/10.1016/j.ejpsy.2025.100290>

Received 30 August 2024; Accepted 9 January 2025

Available online 24 January 2025

0213-6163/© 2025 The Authors. Published by Elsevier España, S.L.U. on behalf of Sociedad Española de Psiquiatría y Salud Mental. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

recent review of interventions to reduce stigma highlighted that few studies have used well adapted and validated outcome measure for stigma, particularly in Low and Middle Income Countries (LMICs).¹² This is important to note given that stigma is strongly influenced by culture, for example in regard to the way in which mental health difficulties are conceptualised, beliefs about causes of these difficulties, and culturally determined values.⁸

Measuring stigma

Stigma has been studied extensively over the past several decades. This has evolved from qualitative research methods to include a range of methodologies, including self-report and behavioural measures of stigmatisation.¹³ One of the key challenges in stigma research relates to the cacophony of approaches to its measurement. Fox et al.¹⁴ conducted a systematic review of studies using mental illness stigma measures between 2004 and 2014. Over 400 different stigma measures were identified, over two-thirds of which had been created for a specific study and had not been systematically psychometrically evaluated. This suggests that the field is at saturation point with regards to the development of new measures. Clearly, there is a need for greater convergence within the field and this should include psychometric evaluation and validation of existing, well-used measures.

From a global perspective, a further issue within the literature is the predominance of studies focusing on Western, English-speaking countries and cultures. Thornicroft et al.¹² conducted a narrative review of anti-stigma intervention research (1970–2012) and found that 83 % of studies took place in high-income countries, with just 17 % taking place in middle-income countries. Strikingly, fewer than 30 % of studies took place in a country other than the US. This indicates a need for research across a wider range of cultural settings, to better understand cross-cultural differences in stigma.¹⁵ Additionally, there is a need for further research within LMICs, given that the generalisation of methods and findings from research conducted in high-income countries is not advisable.¹²

Progression of such research, is, however, a challenge in the context that most stigma measures have been developed in the English language, for use in English-speaking countries.¹⁶ Efforts to measure stigma in non-English speaking countries may either rely on development of a new measure – a potentially time consuming process, or may take an existing measure to be translated, adapted and psychometrically evaluated within the target cultural context. Research suggests that the latter is more common. Indeed, Yang et al.¹⁷ conducted a systematic review of stigma research with non-Western European cultural groups (1990–2012) and found that 77 % ($n = 151$) of included studies used adaptations of existing, Western-developed stigma measures. While this approach may not account for culturally specific aspects of stigma, and makes assumptions about the generalisability of the underlying theory, the translation and use of existing, standardised measures may facilitate comparisons across linguistic and cultural settings.¹⁶

To summarise, it appears that much stigma research has been conducted in high-income Western countries, yet findings are assumed to be universally applicable rather than culturally specific. Further research is required to better understand cross-cultural differences in stigma, and this depends on developing the research base with respect to stigma measurement. Clearly, there is a need for greater convergence within the field of stigma measurement in general, and this should include psychometric evaluation and validation of existing, well-used measures.

The AQ-27

Within Fox et al.'s¹⁴ review, Corrigan et al.'s¹⁸ Attribution Questionnaire was identified as one of the most widely cited stigma measures. To date, the paper has been cited 1830 times on Google Scholar (checked on 10th March 2024). The AQ-27 is a self-report measure of public stigma which was developed in the USA in 2003. It contains a brief

vignette, as follows:

'Harry is a 30-year-old single man with schizophrenia. Sometimes he hears voices and becomes upset. He lives alone in an apartment and works as a clerk at a large law firm. He has been hospitalized six times because of his illness'.

This is followed by twenty-seven statements which measure nine domains related to stigma: blame, anger, pity, help, dangerousness, fear, avoidance, segregation and coercion. Respondents rate their agreement with each statement on a nine-point Likert scale. Higher scores indicate more stigmatising views towards people with mental illness. A short form version of the AQ-27 (the AQ-9)¹⁹ was also developed by the original authors of the measure by selecting the single item that loaded most onto each factor.²⁰

The AQ-27 was originally designed to measure stigma towards people with schizophrenia as the condition is frequently associated with public perceptions of dangerousness.¹⁸ Contemporary research suggests that schizophrenia remains one of the most stigmatised psychiatric diagnoses today.^{21,22} A multinational study by Thornicroft et al.,¹ surveying 732 people with a diagnosis of schizophrenia across 27 countries identified high rates of experienced discrimination, most commonly within friendships, family relationships and in finding and maintaining employment.

Theoretical underpinnings of the AQ-27

The AQ-27 is underpinned by attribution theory, a social cognitive theory which has been applied to understand the relationship between mental health stigma and discriminatory behaviour, in relation to beliefs about causality (personal responsibility for causing one's difficulties) and controllability (the amount of influence an individual can exert over their difficulties).²³ These attributions are thought to lead to differential emotional responses (e.g., pity, anger, fear), which lead to helping or punishing behaviour. The AQ-27 is underpinned by a nine-factor path model which suggests that individuals are more likely to respond negatively to a person with a label of mental illness when they are judged to have a high degree of control over their presentation (e.g., with anger, leading to avoidance and withholding help). Additionally, fear has been found to be a strong predictor of avoidance and support for coercive treatment.¹⁸

Approaches to questionnaire translation and cross-cultural adaptation

It is important to note that questionnaire translation, cross-cultural adaptation and cross-cultural validation are each distinct concepts. We briefly define these terms here. Translation can be defined as the process of transferring meaning from a 'source language' (the primary language in which a measure is written) into a 'target language'.²⁴ This involves consideration of linguistic elements including accuracy, fluency and conceptual equivalence.²⁵ Cross-cultural adaptation considers both language translation and the identification of differences between the 'source culture' and 'target culture' to maintain the equivalence of concepts between both cultural groups. Note that cross-cultural equivalence encapsulates several aspects,¹⁶ including semantic equivalence (equivalence in the meaning of words), experiential equivalence (the relevance of situations or experiences described for the target population) and conceptual equivalence (the validity of the concept described). Lastly, cross-cultural validation aims to ensure that the translated instrument has the same properties as the original instrument.²⁵ Translated measures need to be psychometrically evaluated within the target cultural context.²⁶

Translation and cross-cultural adaptation are complex processes which requires a rigorous, multi-step and collaborative approach. Guiding frameworks have been produced to support the cross-cultural adaptation of self-report measures,¹⁶ such as Beaton et al.'s, 'Guidelines for the Process of Cross-Cultural Adaptation of Self-Report

Measures'.²⁷ Additionally, a variety of translation frameworks are available and these approaches have been reviewed and critiqued extensively within the literature.^{24,25,28} The translation framework used will impact on the quality and validity of the translated measure.

Research questions

The overarching purpose of this systematic review is to review and synthesise the literature in relation to the translation processes of the AQ-27, including assessment of the quality of the translation processes, and associated psychometric properties of translated versions. The review is précised by a broader review of research which has adopted a translated version of the AQ-27 (Part I), followed by a more in-depth review and synthesis of studies which have used a primary translation of the AQ-27 (Part II). Taken together, these components allow us to review the way in which research and literature in the use of the AQ-27 is developing outside English-speaking populations.

Part I: Overview of the use of translated versions of the AQ-27. With what populations, and within what cultural contexts have translated versions of the AQ-27 been used? The purpose of this element is not primarily to establish or summarise the main findings from these papers, but rather to identify the countries and populations in non-English speaking countries in which AQ-27 research is active.

Part II: Assessment of the quality of the translation process, within original translation studies, as well as a review of the psychometric validation of the associated translated version. This component was intended to consider in more detail a smaller subset of papers which had developed a primary translation of the AQ-27 into a different language.

- What languages has the AQ-27 been translated into, from English?
- What is the quality of the procedures used to translate and adapt the AQ-27?
- What is known about the reliability and validity of translated versions of the AQ-27?

Method

Registration

This systematic review was registered on the International Register of Prospective Systematic Reviews (PROSPERO) on 29th June 2023 (registration number CRD42023440611).

Search strategy

The systematic search was completed on 19th September 2023, followed by an update search on 14th January 2024. Searches were carried out with a date limitation from July 2003 until 19th September 2023, in three electronic databases: MEDLINE (PubMed), Web of Science and PsycINFO (EBSCO). To increase the chance of retrieving international papers, Google Translate was used to translate key search terms into the ten most common languages spoken worldwide²⁹ (Mandarin Chinese, Spanish, Hindi, Portuguese, Bengali, Russian, Japanese, Yue Chinese, Vietnamese and Turkish) and these were added to the search strategy. Therefore, the search terms used were:

“attribution questionnaire” OR “AQ-27” OR “AQ27” OR “问卷分配” OR “asignación de cuestionario” OR “প্ৰশ্নাবলী অসাइनমেট” OR “atribuição de questionário” OR “প্ৰশ্নপত্ৰ নিয়োগ” OR “задание анкеты заданије анкеты” OR “アンケートの割り当て” OR “bài tập câu hỏi” OR “anket ödevi”.

Eligibility criteria

Inclusion criteria

Studies included in the review were published, peer-reviewed

empirical studies which used a translated version of the AQ-27 (from English, into another language) to measure stigma, primarily towards people with schizophrenia. Studies which translated an existing abbreviated version of the AQ-27, such as the AQ-9 were included.

For Part II, an additional criterion was applied. Only studies carrying out an original translation of the AQ-27 were included (i.e., studies which used an existing translated version of the measure were excluded).

Exclusion criteria

Studies were excluded based on the following criteria:

- The AQ-27 was explicitly modified to measure stigma towards a condition other than schizophrenia, or stigma towards mental illness in general, however modifications to the wording or structure of the AQ-27 as part of a translation process were included
- The study assessed stigma towards multiple conditions (i.e., the primary focus was not schizophrenia).
- The AQ-27, or abbreviated version was not used in full (e.g., only one subscale was used).
- It was not explicitly stated that the AQ-27 was translated into another language.
- Articles not available in English language.
- For Part II, studies which reported carrying out an original translation, but provided no description of the translation process (as this prohibited any assessment of the quality of the translation process).

We recognised that exclusion criterion (e) is arguably in tension with the core project aims. However, the use of raw machine translation output alone, without the input of qualified human translators, was ruled out for the purposes of the current review due to concerns around the quality and accuracy of the translations. While neural machine translation (NMT), used by systems such as Google Translate is widely regarded as the best performing type of machine translation invented to date, NMT can be inaccurate, is known to output words that do not exist in the target language, and can also amplify biases.³⁰ Moreover, despite literature calling for greater emphasis on publication of non-English papers, the reality remains that most scientific literature is published in the English language, arguably limiting the practical impact of this pragmatic decision.^{31,32}

Screening and selection

Studies identified by the searches were extracted into Microsoft Excel. After duplicates were removed, titles and abstracts were screened for eligibility and removed if they clearly did not meet inclusion criteria. The remaining articles were read in full, and if they were excluded they were coded as to the primary reason for exclusion. Where multiple exclusion criteria applied, the most fundamental exclusion criterion was cited (e.g., studies which did not use the AQ-27, or did not use a translation of the AQ-27). A subset of full-text articles (20 %) were checked by the fourth author, blind to the ratings of the primary reviewer to ensure that they met eligibility criteria.

Quality assessment

The COSMIN Study Design Checklist³³ was used to assess the methodological quality of the translation processes. Additionally, selected items from the COSMIN were used to assess the validity and key psychometric properties of the translated measures. (eTable 1). Each item from the COSMIN is rated on a four-point scale, whereby a score of four indicates the highest methodological quality. Items are weighted according to relative importance. While the COSMIN does not require the use of an overall quality rating, in the present study we calculated a total score by summing the scores for all elements considered. Therefore, the maximum possible overall score was sixty.

Table 1
Overview of Study Characteristics (Part II).

Authors (year)	Version of AQ; number of citing papers in the current review	Country	Study design	Sample size, age range (mean), % female	Participant occupation	Aims	Main findings
Spanish (n = 3)							
Muñoz et al. (2015)	Spanish AQ-27, AQ-27-E; 7 citations	Spain	Translation and psychometric evaluation	439, mean age 39 years, 52.6 % female	Residents in Madrid	To translate and analyse the psychometric properties of the Attribution Questionnaire for use in Spanish-speaking populations (AQ-27-E), and to test the dangerousness and responsibility models of mental illness stigma in a Spanish sample.	"The AQ-27-E has acceptable psychometric properties comparable to previous versions, which can be used to assess stigma in Spanish-speaking populations."
Chamorro Coneo et al. (2022)	Colombian-Spanish adaptation of the AQ-27; 0 citations	Colombia	Cross-sectional	271, 18–79 years (32), mean age 32 years, 67.4 % female	Community sample	To examine pathogen-disgust sensitivity and danger appraisal mechanisms in responses of stigma towards SMI.	"Pathogen avoidance and danger appraisal systems interplay in the generation of discriminatory behaviour towards SMI."
Crespo et al. (2008)	Spanish AQ-27; 0 citations	Spain	Cross-sectional	439, mean age 39 years, 52.6 % female	Community sample from Madrid	To analyse the stigma associated with severe and persistent mental illness in the general population of Madrid.	"Most of the participants showed a helping attitude toward the mentally ill persons, and especially, a disposition to coerce them into treatment."
Chinese (n = 2)							
Chiu et al. (2021)	Modified Chinese AQ; 1 citation	Taiwan	Cross-sectional	123, mean age 21.7 years, 41.5 % female	Medical students	To compare the differences of public stigma, self-stigma, and social distance associated with schizophrenia between old and new name of schizophrenia in Taiwanese medical students.	"After renaming schizophrenia, we noted significant differences in the scores in the modified AQ, the perceived psychiatric stigma scale, and the modified social distance scale in all participants and the fourth-year students, respectively."
Ho et al. (2018)	Chinese translation of AQ-9; 0 citations	Hong Kong	Cross-sectional	218, 17–51 years (22.4), 67 % female	University students	To evaluate the latent profiles of social stigma related to mental illness in the under-researched Chinese context through Factor Mixture Analysis.	"Most of the sample belonged to the low-stigmatizing class, with low to moderate expressions of stigma toward PLMI. The high-stigmatizing class was significantly more likely to be male, not working, and younger and to report significantly higher social distance, personal distress, and empathetic concern."
Italian (n = 1)							
Pingani et al. (2012)	Italian AQ-27, AQ-27-I; 4 citations	Italy	Translation and psychometric evaluation	214, 18–89 years (40.2), 52.3 % female	Relatives of university students	To translate the Attribution Questionnaire-27 (AQ-27) to the Italian language (AQ-27-I), and to examine the reliability and validity of this new Italian version.	"The AQ-27-I demonstrated acceptable internal consistency. Test-retest reliability was also satisfactory. Fit indices of the model supported the factor structure and paths. The AQ-27-I is a reliable measure to assess stigmatizing attitudes in Italian."
Arabic (n = 1)							
Saguem et al. (2021)	Arabic AQ-27; 2 citations	Tunisia	Translation and psychometric evaluation	310, 18–29 years (22.6), 41.9 % female	University students	To translate and validate the AQ in Arabic, by assessing its content validity, construct validity and reliability.	"The Arabic AQ showed acceptable psychometric properties in the assessment of stigma in the Tunisian population. Structural equation models for the responsibility and dangerousness models were mostly supported. The Arabic version of AQ is valid and reliable for the assessment of stigma in Tunisian and Arabic-speaking populations."
Hebrew (n = 1)							
Romem et al. (2008)	Hebrew AQ; 1 citation	Israel	Quasi-experimental	136, mean age 26.1 years, 14.7 % female	Third year nursing students	To evaluate the degree to which a four-week psychiatric clinical clerkship alters nursing	"After the clinical clerkship, students became more compassionate and less

(continued on next page)

Table 1 (continued)

Authors (year)	Version of AQ; number of citing papers in the current review	Country	Study design (pre/post intervention)	Sample size, age range (mean), % female	Participant occupation	Aims	Main findings
Turkish (n = 1)							
Akyurek et al. (2019)	Turkish AQ-27; 0 citations	Turkey	Translation and psychometric evaluation	424, mean age 36.9 years, 52.1 % female	Hospital visitors	To translate the AQ-27 into Turkish and evaluate the reliability and validity of the new Turkish version on a multi-centred selected adult sample.	students' attitudes toward individuals with mental illness. frightened by psychiatric patients, were more willing to care for individuals with mental illness and expressed less need to segregate them from the community. "A good internal consistency was obtained, and a statistically significant test-retest reliability was detected. Fit indices of the model supported the factor structure and paths. AQ-27-T was determined as a reliable and valid questionnaire assessing stigmatization toward mental illness in Turkish population."
Sinhalese (n = 1)							
Baminiwatta et al. (2023)	Sinhalese AQ-9; 0 citations	Sri Lanka	Cross-sectional	405, mean age 39.6 years, 90.6 % female	Nurses	To assess whether higher trait mindfulness among Sri Lankan nurses was linked to lower stigma towards psychiatric patients, and whether compassion mediated this relationship.	"Those with higher trait mindfulness were more likely to believe they would help a person with mental illness, and less likely to believe a person with mental illness should be avoided or segregated from the society. Compassion partially mediated the effects of trait mindfulness on helping and avoidance."
Bengali (n = 1)							
Giasuddin et al. (2015)	26-item Modified Corrigan Attribution Questionnaire (MCAQ); 0 citations	Bangladesh	Cross-sectional	200, mean age of first years 18.9, mean age of fifth years 23.4, 59 % female	First and fifth-year medical students	To explore stigma among medical students toward persons with mental disorders and their attitudes toward psychiatry.	"Upper medical school year, older age, mother's lower academic level, upper and lower socioeconomic level affiliation and self-consultation for mental or neurological complaints were associated with increased stigma toward PMDs. More favourable attitudes toward psychiatry were found in upper medical school year and were significantly associated with female gender and middle socioeconomic level affiliation."
Finnish (n = 1)							
Ihalainen-Tammlander et al. (2016)	Finnish AQ-27; 0 citations	Finland	Cross-sectional	264, mean age 48 years, 98 % female	Nurses in primary healthcare	To describe nurses' attitudes towards people with mental illness and examine factors associated with their attitudes in primary care health centres.	"Nurses' attitudes towards people with mental illness in general were positive in primary care health settings. Younger nurses expressed feeling afraid of mentally ill patients. They not only lacked a feeling of safety around these patients but were also often of the opinion that people with mental illness should be segregated from the general population."

Translation standards outlined within the COSMIN focus on key processes such as completing forward and backward translations, ensuring that the translation is reviewed by a committee and conducting a preliminary pilot study. These processes are critical to achieving linguistic and cross-cultural equivalence and checking the validity of the translated version.²⁷ The COSMIN has been used in a previous systematic review relating to questionnaire translation.³⁴

Using the COSMIN, the first author independently conducted quality

assessments. For inter-rater reliability, the fourth author completed quality ratings for 25 % of included studies ($n = 4$). Any discrepancies were discussed and resolved.

Data extraction

For Part I of the review, the following data were extracted: name of translated measure, language, country, study design, sample size and

demographic information, research aims and main findings.

Part II of the review focused on studies which carried out an original translation of the AQ-27. Information relating to the translation method, and psychometric properties, including factor structure, internal consistency and test-re-test reliability were extracted. This was guided by the COSMIN and informed by quality criteria reported elsewhere.³⁵ Details of any modifications to the AQ-27 were extracted.

Analysis

For Part I, studies and main findings are presented in a table, grouped by country, and key characteristics are summarised narratively. The intention is to allow an overview of the scope of the extant AQ-27 literature within each country. For Part II, a narrative synthesis approach³⁶ was primarily used, combined with visual synthesis of patterns in relation to the quality appraisal (i.e. colour coding) and tabular representation of psychometric properties. Studies were grouped by language and the version of the measure used. Studies were ordered according to frequency of the translation (most translations first) and year of publication (newest first).

Results

Search results

A PRISMA Flow Diagram is shown in Fig. 1. A total of 1404 papers were identified from the initial searches. Following removal of duplicates, 1099 papers remained to be screened. After title and abstract screening, 273 papers were read in full and assessed against the eligibility criteria.

Of note, six papers were excluded due to the full-text articles being published only in a language other than English. These included a German translation of the revised AQ-9, adapted for adolescents,³⁷ an adaptation of the Portuguese version of the AQ-27 for Brazilian speakers,³⁸ an 8-item Spanish translation of the revised AQ-9 for adolescents³⁹ and a Spanish translation of the AQ-14.⁴⁰ It is not known if these papers would have been included in either Part I or Part II had English translations been available. Of the excluded papers, German is the only language which has not been represented within the current review as a result of this exclusion criterion.

Forty-one studies were identified as eligible for inclusion in Part I of the review. Of those, two papers were obtained during the updated search. The 41 papers were then screened for eligibility for inclusion in Part II of the review. Twelve studies were identified as eligible. Of the

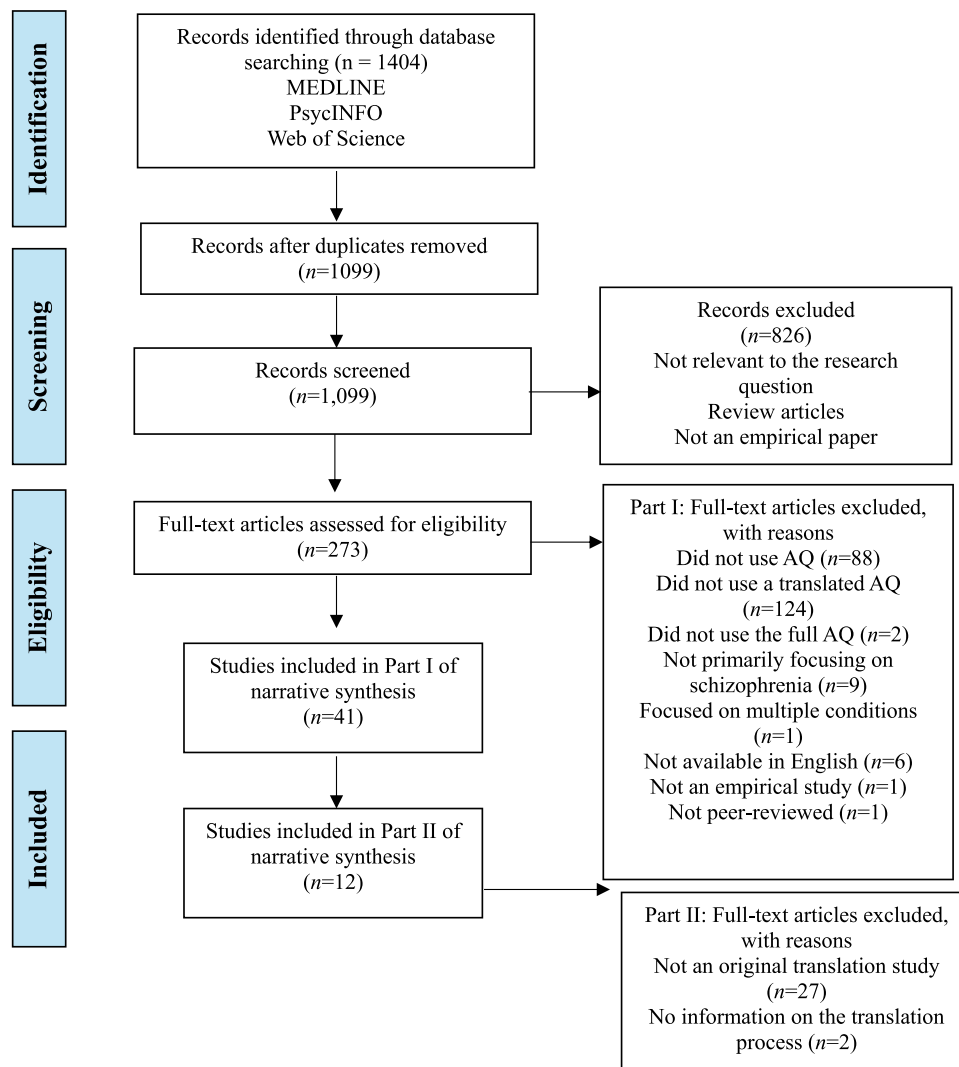


Fig. 1. PRISMA Flow Diagram.⁴¹

papers independently checked by LM there was 100 % agreement.

Part I: Overview of the Use of Translated Versions of the AQ-27: With What Populations, and Within What Cultural Contexts Have Translated Versions of the AQ-27 Been Used?

Study characteristics

Language and country of study

Forty-one studies used a translated version of the AQ-27 to measure stigma towards people with schizophrenia. A summary of the study characteristics and key findings are shown in eTable 2.

We identified that the AQ-27 has been translated into eleven languages, including Spanish ($n = 16$ studies), Portuguese ($n = 6$), Italian ($n = 5$), Chinese languages ($n = 4$; note, the specific Chinese languages were not reported), Arabic ($n = 3$), Hebrew ($n = 2$), French ($n = 1$), Turkish ($n = 1$), Sinhalese ($n = 1$), Bengali ($n = 1$) and Finnish ($n = 1$).

Studies took place across fifteen countries. Most studies took place in Europe ($n = 26$; 63.4 %), with the most common location being Spain ($n = 14$), followed by Portugal ($n = 5$), Italy ($n = 5$), France ($n = 1$) and Finland ($n = 1$). Nine studies (22 %) took place in Asia, including Taiwan ($n = 2$), Hong Kong ($n = 2$), Sri Lanka ($n = 1$), Bangladesh ($n = 1$), Israel ($n = 1$) and Turkey ($n = 1$). Three studies (7.3 %) took place in South America, including Chile ($n = 1$), Colombia ($n = 1$) and Brazil ($n = 1$). Three studies (7.3 %) were carried out in Africa, in Tunisia ($n = 3$).

The total sample sizes for each country represented in the review are shown in Fig. 2. The largest total samples were obtained from Spain ($n = 2597$), Italy ($n = 1379$) and Portugal ($n = 703$).

Participant characteristics

In total, 8709 participants were recruited. Sample sizes ranged from 22,⁴² to 2746.⁴³ Most studies ($n = 35$, 85.4 %) consisted of a majority female sample (≥ 50 %). The mean age of participants, where reported ranged from 17.8 to 54.9 years. Studies sampled from a range of populations, including university students ($n = 17$, 41.5 %), the general public ($n = 8$, 19.5 %), mixed populations ($n = 5$, 12.2 %), health professionals ($n = 4$, 9.8 %), high school students ($n = 3$, 7.3 %), service users ($n = 1$, 2.4 %), service users' relatives ($n = 1$, 1.2 %), school staff ($n = 1$, 1.2 %) and college students ($n = 1$, 1.2 %).

Study design

A wide variety of study designs were observed. These included cross-sectional studies ($n = 17$, 41.5 %), quasi-experimental designs ($n = 8$, 19.5 %), studies investigating measurement properties of the AQ-27 ($n = 7$, 17.1 %), correlational studies ($n = 6$, 14.6 %), randomised controlled trials ($n = 3$, 7.3 %), and mixed designs ($n = 1$, 2.4 %).

Part II: Translations of the AQ-27

Assessment of the Quality of the Translation and Adaptation Process, Within Original Translation Studies.

a) What Languages has the AQ-27 Been Translated Into, From English?

Part II of the review focused on a subset of the studies included in Part I, which reported carrying out an original translation of the AQ-27 (i.e., rather than using an existing translation).

Of the 41 studies initially identified, 14 studies produced an original translation. However, two studies^{44,45} provided no information about the translation process and were therefore excluded. This left 12 studies remaining for inclusion in Part II of the review. Table 1 provides an overview of the study characteristics.

Language and country of study

The 12 original translation studies spanned nine languages, including Spanish,⁴⁴⁻⁴⁶ Chinese languages,^{47,48} and Italian,⁴⁹ Arabic,⁵⁰ Hebrew,⁵¹ Turkish,⁵² Sinhalese,⁵³ Bengali,⁵⁴ and Finnish.⁵⁵ The Spanish AQ-27,⁴⁴ had the highest number of citing papers within the current systematic review ($n = 7$ citations), followed by the Italian AQ-27,⁴⁹ ($n = 4$), Arabic AQ,⁵⁰ ($n = 2$), Chinese AQ⁴⁷ ($n = 1$) and Hebrew AQ-27,⁵¹

($n = 1$). This suggests that the Spanish, Italian and Arabic versions of the AQ-27 are gaining traction.

Studies took place across Asia (Taiwan,⁴⁷ China,⁴⁸ Israel,⁵¹ Turkey,⁵² Sri Lanka,⁵³ Bangladesh⁵⁴), Europe (Spain,^{44,46} Italy,⁴⁹ Finland⁵⁵), Africa (Tunisia⁵⁰) and South America (Colombia⁴⁵).

Participant characteristics

Across the studies, 3004 participants were recruited. Sample sizes ranged from 123,⁴⁷ to 439,⁴⁴ Studies sampled university students^{47,48,50,51} ($n = 5$), the public^{44-46,49-52} ($n = 5$) and nurses^{53,55} ($n = 2$). Most studies ($n = 9$, 75 %) contained predominantly female samples (≥ 50 %). The mean age of participants ranged from 18.9 years,⁵⁴ to 48 years.⁵⁵

Study designs

Importantly, there was significant heterogeneity with regards to the study designs and aims. Only four studies (33.3 %) had a primary aim of translating and psychometrically evaluating the AQ-27; those were the Spanish AQ-27,⁴⁴ Italian AQ-27,⁴⁹ Arabic AQ⁵⁰ and Turkish AQ-27.⁵² The remaining studies consisted of cross-sectional designs^{45-48,53-55} ($n = 7$) and pre/post intervention designs ($n = 1$).⁵¹

b) What is the Quality of the Procedures Used to Translate and Adapt the AQ-27?

Quality Assessment of the Translation Process

Selected items from the COSMIN Study Design Checklist (eTable 1) were used to assess the quality of the translation method. This informed Research Question II(b). Table 2 provides an overview of the findings and full results are provided in eTable 3.

Overall quality ratings varied widely from 25,⁴⁷ to 54,⁵² out of a maximum of 60. The Turkish AQ-27,⁵² was the highest rated translation, followed by the Italian AQ-27,⁴⁹ and Arabic AQ,⁵⁰ scoring 48 and 44, respectively. All of these studies were primarily focused on translation and psychometric evaluation of the AQ-27. However, two-thirds of the translation studies ($n = 8$) were not focused on translation of the AQ-27 as a research aim and subsequently provided limited information about the translation method or framework. This significantly limited our ability to appraise the quality of the translation approach.

Nonetheless, the quality appraisal highlighted some key themes. Firstly, in the COSMIN (and indeed, in most translation guidelines²⁵) it is advised that at least two forward and backward translations are completed by independent translators, to enable the translations to be synthesised and for any differences to be resolved. In the current review, most studies ($n = 10$, 83.3 %) had completed at least one forward and one backward translation, but only four studies^{49,50,52,55} (33.3 %) had completed multiple forward and backward translations. A key limitation of this simple 'direct and back' method, particularly where only two translations are produced overall, include that this method may focus only on linguistic equivalence while neglecting cultural considerations.²⁴

Questionnaire translation is a complex process which requires a combination of linguistic, cultural and subject matter expertise. As such, it is recommended that forward and backward translators have specific linguistic backgrounds and knowledge.^{27,33} In the current review, many studies did not report on the profiles of the translators, and three studies did not use professional translators at all, but rather, took an 'ad hoc' approach. This included the Spanish AQ-27-E, which was the most widely adopted version within the review. While one could speculate about the possible reasons for this (e.g., lack of time, access to professional translators), this approach is not considered sufficient to produce an accurate and equivalent translation.

A third step which is crucial to the translation process involves carrying out an expert committee review, to consolidate all versions of the questionnaire prior to pilot testing. It is recommended that the multi-disciplinary committee should comprise all translators, and language, culture and subject matter experts, ideally including the original developers of the measure.²⁷ This 'team-based' approach is considered

essential to establishing cross-cultural equivalence.²⁴ In the current review over half of the studies ($n = 7, 58.3\%$) did not involve an expert committee in the translation process.

The final step of questionnaire translation is to carry out pilot testing within the target setting.²⁷ The purpose of this is to check respondents' understanding of the questionnaire items. Within the review, half of the included studies ($n = 6, 50\%$) did not carry out pilot testing.

While the current systematic review focused on approaches to translation, rather than cross-cultural adaptation, it was interesting to note that one only study (the Turkish AQ-27)⁵² referred to cultural

adaptation. Akyurek et al.⁵² describe in detail a multi-step adaptation method, citing Beaton et al.'s²⁷ widely cited cross-cultural adaptation guidelines. This facilitated auditing of the translation methodology and provides increased assurance of the quality and cross-cultural equivalence of the measure.

c) What is Known About the Reliability and Validity of Translated Versions of the AQ-27?

Data were extracted relating to the reliability and validity of the translated measures, where provided. Results are shown in Table 3.

Table 2
Overview of the Quality of Translation Processes.

Authors (year)	Name of measure	Country	Translators	Forward-backward translation	Expertise, and mother tongue of translators ^b	Independent translation	Translation is reviewed by a committee	Pilot testing to check the comprehensibility of the instructions, items and response options	Overall quality rating (max = 60) and ranking based on score
Spanish (n=3)									
Muñoz et al. (2015) ^a	Spanish AQ-27 (AQ-27-E)	Spain	Study authors.	Yes – one forward translation and one backward translation.	Not reported.	Unclear whether translators worked independently.	No expert committee review.	Yes – “A pilot study was performed with general population (N=20) to study the behavior of the items, their comprehension, and adequacy.”	38, 5 th
Chamorro-Coneo et al. (2022)	Colombian -Spanish adaptation of AQ-27	Colombia	Two professional translators.	Yes – one forward translation and one backward translation.	Not reported.	Assumable that translators worked independently.	No expert committee review.	Yes – “The Colombian-Spanish AQ was pilot tested (N = 21) with the other variables of this study. Results showed good comprehension of the items amongst participants and adequate distribution of the scales' scores.”	35, 6 th
Crespo et al. (2008)	Spanish AQ-27	Spain	Study authors.	Yes – one forward translation and one backward translation.	Not reported.	Translators were not independent.	No expert committee review.	No pilot testing.	30, 10 th
Chinese (n=2)									
Chiu et al. (2021)	Modified Chinese AQ	Taiwan	Not reported.	No –forward translation only.	Not reported.	Not reported.	No expert committee review.	No pilot testing.	25, 11 th
Ho et al. (2018)	Chinese AQ-9	Hong Kong	Study authors.	Yes – one forward translation and one backward translation.	Not reported.	Translators were not independent.	Yes – the Chinese AQ-9 was reviewed by the original authors of the AQ-27.	No pilot testing.	33, 8 th
Italian (n=1)									
Pingani et al. (2012) ^a	Italian AQ-27 (AQ-27-I)	Italy	Three native Italian speakers (forward), one professional translator (backward).	Yes – multiple forward translations and one backward translation.	Forward translators were native Italian speakers, as per guidelines. Backward translators: expertise and mother tongue not described.	Translators worked independently.	No expert committee review.	Yes – “The draft Italian version was tested on 30 Italian-speaking students. A group discussion followed, with students required to answer two questions for each item: “What does this statement mean to you?” and “Is there any other wording that enables this meaning to be expressed more clearly?”	48, 2 nd
Arabic (n=1)									
Saguem et al. (2021) ^a	Arabic AQ	Tunisia	Four professional translators.	Yes – multiple forward and multiple backward translations.	Forward translators: expertise and mother tongue not described. Backward translators: both had no knowledge of the original AQ-27, as per guidelines.	Translators worked independently.	Yes - sixteen experts (adult and child psychiatrists) were asked to judge the clarity, the relevance, the discernment of each item and response validity.	“The Arabic version was tested on 101 medical students. They were asked to comment on each item of the instrument and to mention any other wording that enables its meaning to be	44, 3 rd

(continued on next page)

Table 2 (continued)

				Mother tongue unknown.			expressed more clearly."		
Hebrew (n=1)									
Romem et al. (2008)	Hebrew AQ	Israel	"Three independent judges"	No – forward translation only.	Not reported.	Translators worked independently.	No expert committee review.	No pilot testing.	32, 9 th
Turkish (n=1)									
Akyurek et al. (2019) ^a	Turkish AQ-27	Turkey	Four professional translators, and two native English speakers	Yes – multiple forward translations and multiple backward translations.	Forward: two translators with a medical background and two without a medical background, as per guidelines. Backward: two native English speakers who had no medical background, as per guidelines.	Translators worked independently	Yes – "an expert committee examined all translations, reports and the original version of the survey and developed the penultimate pilot version."	First pre-test: "30 participants completed the Turkish AQ-27 to determine whether the measure was intelligible and appropriate for Turkish culture. Participants were interviewed individually about the lucidity of the items, the accuracy of the reflection of general stigma in the survey, and the usefulness and length of the survey." Second pre-test: "Following cultural adaptations and changes to wording, the Turkish AQ-27 was administered to 52 participants."	54, 1 st
Sinhalese (n=1)									
Baminiwatta et al. (2023)	Sinhalese AQ-9	Sri Lanka	"Bilingual experts"	Yes – one forward and one backward translation.	Not reported.	Unclear whether translators worked independently.	Yes – "revisions by an expert panel".	No pilot testing	34, 7 th
Bengali (n=1)									
Giasuddin et al. (2015)	Bengali 26-item Modified Corrigan Attribution Questionnaire (MCAQ)	Bangladesh	"Expert translators"	Yes – multiple forward and multiple backward translations.	Not reported.	Assumable that translators worked independently.	No expert committee review.	No pilot testing	35, 6 th
Finnish (n=1)									
Ihalainen-Tamlander et al. (2016)	Finnish AQ-27	Finland	Study authors and one professional translator.	Yes – multiple forward and multiple backward translations.	Not reported.	Translators worked independently.	Yes – "the developer of the questionnaire inspected whether the translated version of the instrument corresponded with the original instrument."	"The Finnish instrument was pilot tested in two psychiatric wards with 20 volunteer nurses (not participating in the main study) to find out the clarity and understandability of the items, effectiveness of instructions and time required to complete the questionnaire."	43, 4 th

Note. Colour coding reflects scoring from the quality assessment using the adapted COSMIN Study Design Checklist (0–4). Dark green=4 (very good), light green=3 (adequate), light orange=2 (doubtful), dark orange=1 (inadequate), grey=0 (not reported).

^aStudies with a primary aim of translating and analysing the psychometric properties of the AQ-27.

Reliability

i) Internal Consistency

Internal consistency reflects the extent to which items in a questionnaire, or its subscales are correlated and therefore measure the same construct.³⁵ Cronbach’s alpha (α , expressed as a number between 0 and 1) is a commonly used measure of internal consistency. Alpha values of between 0.7 and 0.95 can be considered indicative of good internal consistency.³⁵

Eight studies (66.7 %) reported on internal consistency for the AQ-27 as a whole and all reported values were above the threshold for acceptability. Subscale alpha values were provided for the Spanish,⁴⁴ Italian⁴⁹ and Hebrew⁵¹ translations (41.7 %, $n = 5$). Low alpha values were reported for the Responsibility ($\alpha=0.39 - 0.615$),^{44,51} Pity ($\alpha=0.494 - 0.676$),^{44,45,49} and Anger subscales ($\alpha=0.521 - 0.577$) across several studies,^{44,49} which may indicate that some subscale items need to be revised or removed. This could be further explored by assessing the

extent to which subscale items correlate with each other and with the total score.⁵⁶ Internal consistency was not assessed for the Finnish AQ-27,⁵⁵ or Sinhalese AQ-9.⁵³

ii) Test-Retest Reliability

Test-retest reliability refers to the degree to which repeated measurements with the same participants under the same conditions produces consistent results.³⁵ The Intraclass Correlation (ICC) is a widely used measure of test-retest reliability.⁵⁷ Values range from 0 to 1, with values closer to 1 indicating stronger reliability.

Only two studies^{49,52} (16.7 %) reported on test-reliability. For the Italian AQ-27,⁴⁹ both total and subscale ICCs were provided. The total ICC (0.72) was within the range for moderate reliability⁵⁷ (0.5–0.75) and subscale ICC values ranged from 0.51 (moderate) for Anger, to 0.89 for Fear (approaching excellent reliability). For the Turkish AQ-27,⁵² both total and item Pearson correlation coefficients were provided as a measure of test-retest reliability. The total Pearson correlation

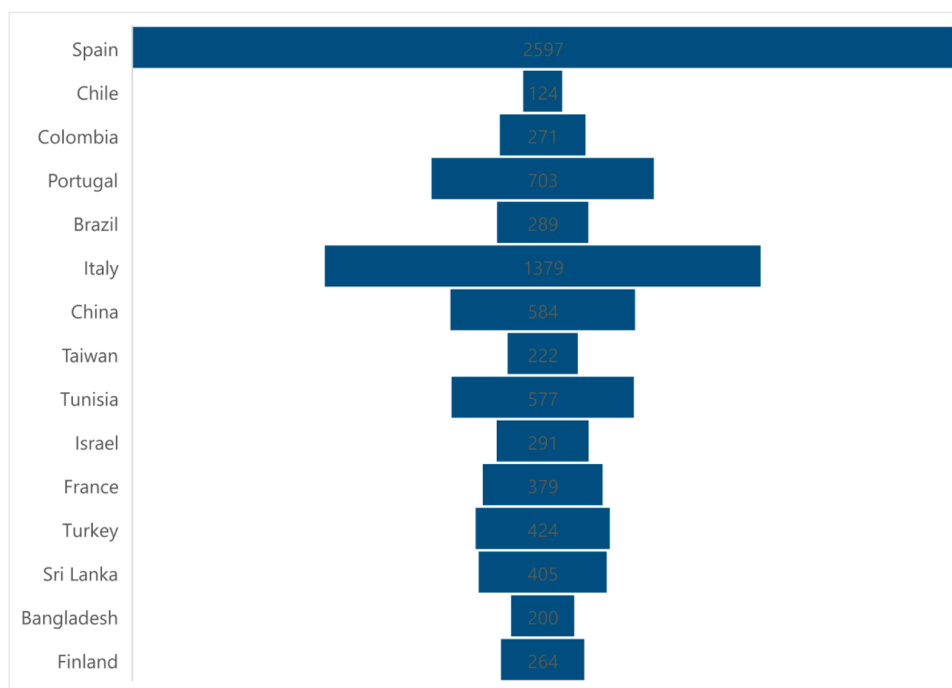


Fig. 2. Total Sample Size for Each Country Represented Within the Review.

coefficient (0.793) suggested that the Turkish AQ-27 had adequate test-retest reliability.⁵⁸

Validity

i) Factor Structure (Structural Validity)

Factor analysis explores the relationship between questionnaire items and underlying dimensions of the measured construct (i.e., factor structure) which may explain these relationships.⁵⁹ The two main forms of factor analysis are Exploratory Factor Analysis (EFA), which explores the underlying relationships between variables, and Confirmatory Factor Analyses (CFA), which assesses whether the data fit a hypothesised measurement model. The AQ-27 was originally conceptualised as consisting of a nine-factor structure.¹⁸

In the current review, two-thirds of the included studies ($n = 8$, 66.7 %) did not carry out a factor analysis. CFA was carried out for the Italian⁴⁹ and Turkish AQ-27,⁵² and in both cases, results supported the original nine-factor structure of the AQ-27. EFA was carried out for the 20-item, Modified Chinese AQ,⁴⁷ resulting in a six-factor solution. The Arabic AQ⁵⁰ was derived by translating an existing 21-item version of the AQ, and consists of a seven-factor structure. The 21-item measure excluded the Segregation subscale (items 6, 15 and 17) and Coercion subscale (items 5, 14 and 25) due to a lack of support for these subscales in previous translated versions.⁴⁶

Discussion

Since its inception in 2003, the AQ-27 has become a well-established measure of public mental illness stigma in the English language. This was the first systematic review to explore the use of translated (non-English language) versions of the AQ-27 to measure stigma towards people with schizophrenia. In Part I, we conducted a review of studies which had used a translated version of the AQ-27 in pursuit of a wider research question, and in Part II we considered in more detail the studies which had conducted a primary translation of the AQ-27. The methodological quality of the translation processes was assessed using COSMIN criteria,³³ and psychometric data were reviewed.

Part I of the review identified that to date, the AQ-27 has been translated into eleven languages and implemented across fifteen

countries. As highlighted in eTable 2, it has been used in a wide range of studies considering a range of different research questions and adopting different methodologies with a range of different types of samples (see also Fig. 2). There are few obvious findings from these studies which can be synthesized, except it is clear that the AQ-27 appears to be being used in a diverse range of ways including determination of between-group differences, assessment of potential outcomes from interventions, and as an independent variable in a range of different ways.

Regarding geographical distribution, Western Europe was grossly over-represented in the review. Most studies (63.4 %) took place in Europe, with the largest samples being obtained from Spain, Portugal and Italy. In particular, the Spanish literature (predominantly arising from Spain) appears relatively well advanced, which is possibly related to the fact that three separate efforts appear to have been made to develop a translated AQ-27 in Spanish (a fact that is not without its problems, considered in more detail within Part II).

Outside of Europe, a smaller proportion of studies took place in Asia (22 %), Africa (7.3 %) and South America (7.3 %). Similar findings were reported in a previous review (1990–2012) by Yang et al.¹⁷ The current review therefore adds to existing literature which suggests that stigma research is overall skewed towards ‘WEIRD’ countries and populations. An important implication of this is the need to avoid making assumptions about the suitability of the AQ-27 in contexts in which stigma research is less well established, i.e. many LMICs (Low- and Middle-Income Countries). In particular, it is noted that the underlying assumptions of wider Attribution Theory – on which the AQ-27 significantly draws – may not necessarily generalise into other cultures directly. One recommendation therefore is that future research considering cross-cultural translation and adaptation of instruments such as the AQ-27 should ideally take a more ‘bottom up’ approach where the underlying theory behind the measure is first developed and adapted in the relevant cultural context before the translation process begins. For researchers considering adopting the AQ-27 directly in a non-English context, consideration should be given to cultural equivalence of the relevant underlying theoretical concepts. Additionally, factor analysis is required following development of a translated measure in order to establish the underlying factor structure.

Part II of the review considered, in more detail, the studies which had

Table 3
Reliability and Validity of Translated Versions of the AQ-27.

Authors (year)	Name of measure, location	Participant occupation, sample size, age range (mean), % female	Modifications to items	Modifications to vignette	Changes to factor structure, factor analysis (e.g. CFA, EFA)	Internal consistency (Cronbach's alpha)	Test-retest reliability (e.g. intraclass correlation coefficient)
Spanish (n = 3)							
Muñoz et al. (2015) ^a	Spanish AQ-27, AQ-27-E, Spain	Residents in Madrid; 439, mean age 39.01 years, 52.6 % female	No changes- retained 27-item AQ.	No changes – “AQ-27 includes a neutral vignette that represents a hypothetical person (Harry) who suffers from a severe mental illness.”	No changes - retained the original nine factor structure. No factor analysis.	Total= 0.855 Fear = 0.896; Anger = 0.577 Help = 0.766; Dangerousness = 0.849; Avoidance = 0.730; Segregation = 0.848; Pity = 0.494; Responsibility = 0.390; Coercion = 0.478 Total alpha not reported.	Not reported.
Chamorro Coneo et al. (2022)	Colombian-Spanish adaptation of AQ-27, Colombia	Community sample; 271, 18–79 years (32), mean age 32 years, 67.37 % female	Reduced the number of items to 20, however the process by which this was achieved is not described	No changes – “The AQ-27 in Colombian Spanish comprised four vignettes describing the story of “Juan”, a man with a SMI. The story in each vignette was different regarding Juan’s aggressiveness and causes associated with the cause and exacerbation of his symptoms.”	Factor structure unclear. No factor analysis.	Total alpha not reported. Anger = 0.81; Fear = 0.96; Helping/avoidance = 0.84; Coercion/ segregation = 0.86; Responsibility = 0.60; Pity = 0.55	Not reported.
Crespo et al. (2008) ^a	Spanish AQ-27, Spain	Community sample; 439, mean age 39.01 years, 52.6 % female	No changes- retained 27-item AQ.	No changes – used neutral version of the vignette.	No changes - retained the original nine factor structure. No factor analysis.	Total = 0.76 Subscale alphas not reported	Not reported.
Chinese (n = 2)							
Chiu et al. (2021)	Modified Chinese AQ (20 items), Taiwan	Medical students; 123, mean age 21.7 years, 41.5 % female	“Due to the similarity after translation into Chinese, we extracted 20 items of the Corrigan’s attribution questionnaire according to experts’ opinions for this study” - removed items 4, 12, 19, 21, 22, 24 and 26	Modified the vignette to compare the old and new name of schizophrenia in Taiwan (“disorder with dysfunction in thought and perception”).	Items were grouped into nine subscales. Exploratory factor analysis yielded a six-factor solution.	Total (old name)= 0.83 Total (new name) = 0.82 Subscale alphas not reported	Not reported.
Ho et al. (2018)	Chinese AQ-9, Hong Kong	University students; 218, 17–51 years (22.4), 67 % female	No changes - retained 9-item AQ.	No changes – “John is a single man who lives alone in an apartment and works as a clerk at a large law firm. He was diagnosed with schizophrenia. He often hears voices of unknown origin and becomes upset. He has been hospitalized for two months because of his illness”.	“Preliminary factor mixture analysis supported a one-factor structure for the scale.”	Total = 0.80 Subscale alphas not reported	Not reported.
Italian (n = 1)							
Pingani et al. (2012) ^a	Italian AQ-27 (AQ-27-I), Italy	Relatives of university students; 214, 18–89 years (40.15), 52.3 % female	No changes- retained 27-item AQ.	No changes – “the vignette described ‘Harry’, a 30-year-old single man with schizophrenia”.	Confirmatory factor analysis (CFA) “Our major goal was to determine whether the Italian model mirrored the American; fit indicators were equivalent on the matter”.	Total=0.818 Responsibility = 0.615; Pity = 0.676; Anger = 0.521 Dangerousness = 0.755 Fear = 0.912; Help = 0.814 Coercion = 0.570; Segregation = 0.801; Avoidance = 0.570	Total intraclass coefficient (test-retest reliability) = 0.72 Subscale ICCs ranged from 0.51 (Anger) to 0.89 (Fear)

(continued on next page)

Table 3 (continued)

Authors (year)	Name of measure, location	Participant occupation, sample size, age range (mean), % female	Modifications to items	Modifications to vignette	Changes to factor structure, factor analysis (e.g. CFA, EFA)	Internal consistency (Cronbach's alpha)	Test-retest reliability (e.g. intraclass correlation coefficient)
Arabic (n = 1)							
Saguem et al. (2021) ^a	Arabic AQ, Tunisia	University students; 310, 18–29 years (22.6), 41.9 % female	Translated a 21-item version of the AQ which omitted terms for segregation and coercion.	No changes reported – “The questionnaire starts with a short statement about “Harry,” a 30-year-old single man who works as a clerk in a law firm and who has been hospitalized for schizophrenia.”	Describe a seven-factor model for the 21-item Arabic translation; Responsibility, Pity, Help, Avoidance, Dangerousness, Fear, Anger. No factor analysis.	Total = 0.71 Responsibility = 0.78 Pity = 0.82; Help = 0.72 Avoidance = 0.72 Dangerousness = 0.78 Anger = 0.73; Fear = 0.74	Not reported.
Hebrew (n = 1)							
Romem et al. (2008)	Hebrew AQ, Israel	Third year nursing students; 136, mean age 26.1 years, 14.7 % female	“One statement was excluded due to difficulties retaining the original meaning following translation into Hebrew..”	No changes – “the final questionnaire included vignettes about four 30-year-old men with schizophrenia, which vary in the level of danger and controllability attributed to the patient”.	Six constructs, with 3–4 items each; Responsibility, Pity, Anger, Fear, Willingness to Help, Segregation. No factor analysis.	Total alpha not reported. Subscales (pre/post intervention): Responsibility 0.55, 0.86 Pity = 0.87, 0.83; Anger = 0.87, 0.83; Fear = 0.87, 0.82; Willingness to Help = 0.78, 0.80; Segregation = 0.84, 0.87	Not reported.
Turkish (n = 1)							
Akyurek et al. (2019) ^a	Turkish AQ-27, Turkey	Hospital visitors; 424, mean age 36.9 years, 52.1 % female	“The wording of items 4, 11, 12, 13, 14, 17, 19, 20, 22, 24, 27 were amended to preserve the original meaning, as part of the cultural adaptation process.” - all wording changes are described in full.	No changes–“Hasan is a 30-year-old single man with schizophrenia. Sometimes he hears voices and becomes upset. He lives alone in an apartment and works as a clerk at a large law firm. He had been hospitalized six times because of his illness.”	CFA indicated that the original nine factor structure was supported.	Total = 0.88 Individual items ranged from 0.866 to 0.892	Pearson correlation coefficient (for total score)=0.793 Item correlation coefficients ranged from 0.35 to 0.77
Sinhalese (n = 1)							
Baminwatta et al. (2023)	Sinhalese AQ-9, Sri Lanka	Nurses; 405, mean age 39.6 years, 90.6 % female	No changes - retained 9-item AQ.	No changes – “hypothetical vignette about a man named Harry who has schizophrenia”.	N/A – “each domain in the AQ-9 was measured by only a single item”.	N/A – “each domain in the AQ-9 was measured by only a single item”.	Not reported.
Bengali (n = 1)							
Giasuddin et al. (2015)	Bengali 26-item Modified Corrigan Attribution Questionnaire (MCAQ), Bangla-desh	First and fifth-year medical students; 200, mean age of first years 18.9, mean age of fifth years 23.4, 59 % female	“One question from the original questionnaire was deleted: ‘If I were in charge of the treatment of Hasib, I would force him to live in a group home’, since this service option is unavailable in the country”.	No changes–“The MCAQ provides a brief vignette about Hasib, a 30-year-old single man with schizophrenia who lives alone and works as a clerk at a large private firm. He had been hospitalized six times because of his illness.”	No factor analysis.	Total = 0.71	Not reported.
Finnish (n = 1)							
Ihalainen-Tamlander et al. (2016)	Finnish AQ-27, Finland	Nurses; 264, mean age 48 years, 98 % female	No changes- retained 27-item AQ.	No changes – “Harry is a 30-year-old single man with schizophrenia. Sometimes he hears voices and becomes upset. He lives alone in an apartment and works as a clerk at a large law firm. He has been hospitalized six times because of his illness”.	No changes - retained the original nine factor structure. No factor analysis.	Cronbach's alpha not reported.	Not reported.

^a Studies with a primary aim of translating and validating the AQ-27.

conducted a primary translation of the AQ-27 from English into another language. Overall, these studies can be grouped into a smaller group ($n = 4$) which were primarily focused on translation and validation of the measure,^{44,49,50,52} and another group ($n = 8$) where the translation had occurred in the context of a separate research question. The first group of papers appeared to have notably better rigour and quality of translation methodology. Notably, the rigour and quality of translation methodology did not necessarily appear to correlate with the extent of research activity; the Spanish and Chinese papers are a case in point: these were the only languages where more than one author had approached development of a primary translation, but there were (relative) gaps and important areas for improvement.

Overall, the Turkish,⁵² Arabic⁵⁰ and Italian⁴⁹ versions were rated highest in terms of the quality of the translation processes. While the current systematic review focused on efforts to translate (rather than culturally adapt) the AQ-27, it was interesting to note that Akyurek et al.⁵² were the only authors to address cultural considerations as part of the translation and adaptation process. Future researchers wishing to adapt the AQ-27 for non-English-speaking cultures should consider using translation frameworks which incorporate cultural considerations, as this may increase the validity of the AQ-27 as a measure of mental illness stigma within the target culture. Attribution theory is likely to be implicated in cross-cultural adaptation (e.g., the extent to which respondents view mental distress as being controllable and within one's personal responsibility) and this should be considered as part of the translation and adaptation process. Akyurek et al.'s paper provides an example of how this might be achieved using Beaton et al.'s²⁷ cross-cultural adaptation guidelines. Additionally, researchers should be aware that translation is not equivalent to cross-cultural adaptation and therefore these terms should not be used interchangeably.²⁵ More widely, we hope that our approach to quality appraisal can help authors seeking to develop translated measures to identify important methodological priorities (including for instance pilot testing and use of committees), as well as what information to report in their manuscript.

Unfortunately, these better examples of translation can be contrasted with the majority of the other translation studies, and the review has overall identified many areas in which translation processes were weak or where insufficient information was provided to make a judgement. For instance, several studies appeared to adopt a relatively crude forward-backward translation approach, without committee involvement. It has been argued that forward-backward translation should not be relied upon exclusively as a means of producing an equivalent translation, since this may overemphasise linguistic equivalence while neglecting to account for cultural variation and idiosyncrasies.⁶⁰ Consensus within the field is that forward-backward translation should be combined with a committee or team-based approach.²⁴ As stated by Behr⁶⁰:

A methods description along the lines of 'We translated and back translated the questionnaire to check for equivalence,' which is all too common, should not be regarded as sufficient evidence of a flawless and equivalent translation. Efforts should be directed towards ensuring quality in the translation itself – by committee or team approaches; by the involvement of suitable translation, content, and survey experts; and by thorough documentation of the translation process, including problems and intentional deviations from a source questionnaire. (Behr, 2017, p. 582)

This is reflected within cross-cultural adaptation guidelines^{16,27} and quality criteria³³ which recommend that translations are reviewed by an expert committee and then pilot tested within the target cultural context. However, within the current review, over half of the included studies (58.3 %) did not involve an expert committee and half did not carry out pilot testing. Furthermore, three studies did not use professional translators. This may have implications for the quality of the data obtained using these translated measures.^{24,60}

Beyond this specific point, there are many more pieces of important

information which translation studies should calculate and report. Very few studies provided information regarding the profiles and expertise of the translators, and most studies did not refer to any standardised translation protocol. Questionnaire translation guidelines^{16,27} emphasise the importance of fully documenting each step of the translation process, to enable the quality of the translation approach to be evaluated. Whilst this may be a reflection on overall research quality, an alternative reason for failure to include these components may be authors' concerns about adding to the length of their journal articles; authors should thus be encouraged to include such material as supplementary material or publish such material in relevant 'open' repositories. Such practices make comparative assessment of quality much easier, and allows the literature to much more effectively build on what has gone before.

Following translation of a measure, it is important to assess its psychometric properties in the translated language.^{17,27} Again, this is an area where translation studies show significant potential for improvement, and where future authors would be strongly encouraged to exert efforts. Whilst Cronbach's alpha was reported frequently (though even here, four of the studies did not report this data at all) only four studies carried out a factor analysis, and only two studies reported on test-retest reliability. The findings suggesting poor reliability of translated versions of the AQ-27 at a subscale level warrants further research.

Beyond the limitations observed in the synthesised data, it is also important to briefly reflect on the limitations inherent in the review methodology. Arguably the largest limitation is that for pragmatic reasons, non-English publications were excluded from the systematic review. If resources had not been constrained, we would have ideally developed a research team that would have allowed inclusion of papers in all of these languages. Whilst there is some evidence to suggest that excluding non-English papers from systematic reviews may have minimal impact (since most scientific papers are published in the English language),⁶¹ we did identify six articles which were not possible to include because they lacked an English translation. This does suggest that future reviews of translated measures may be improve at least modestly if attention is given to processes to support the inclusion of non-English language papers, including where necessary international collaborative efforts and better inclusion of native speakers or translators.

We deliberately only sought peer-reviewed, published studies as we aimed to identify translated versions of the AQ-27 which were likely to be of a sufficient quality to be of value to future researchers. However, it is possible that the exclusion of grey literature reduced the comprehensiveness of the review. This may be an important consideration for future systematic reviews (e.g., given concerns about Western-centred biases in academic publishing).⁶²

Conclusion

This systematic review provides an overview of the use of translated versions of the AQ-27, and an assessment of the methodological quality of the translation approaches. Some relatively robust translation approaches were identified (e.g., for the Turkish,⁵² Arabic⁵⁰ and Italian⁴⁹ adaptations), but more widely there was significant scope for improvement in the quality of translation approaches or at least better reporting of quality markers in published studies. We hope that the approach to consideration of quality provides a framework on which future researchers can build, and allows a reduction in duplication of research efforts. A stepwise and incremental approach to stigma research is important to reduce the likelihood of replicating the cacophonous situation in relation to stigma measures that exists in the English-speaking world.

For most translated versions, therefore, researchers should avoid making assumptions about the quality of the original translation methodology used to develop existing measures before adopting them. A poor-quality translation could potentially invalidate conclusions drawn

from the data.²⁴ This is particularly important in light of the wider research situation involving use of the AQ-27 in non-English-speaking regions; whilst eTable 2 highlights a relatively broad range of research activity, particularly in some regions, it is a concern that the underpinning translations of the AQ-27 leave room for improvement in several ways. The research situation in Spain (and in Spanish versions more widely) is arguably a particular case in point, where research activity is most advanced, but where three translations of the AQ-27 exist, all of which appear to have room for improvement.

In future, researchers wishing to develop their own translations of the AQ-27 should be aware that a systematic and rigorous approach, based on a robust translation framework and ideally involving a committee approach is recommended to ensure that the translated measure is valid and equivalent within the target culture.²⁴ A variety of translation frameworks,^{24,27} and quality appraisal tools are available to support this.³³ Attention should also be given to culturally inappropriate assumptions which are inherent in any underlying theory.

Considering the context much more broadly, one must remember that stigma is itself a social and cultural construction.^{6,63} When considering the cross-cultural adaption of existing stigma measures, it is important to note that many tools, including the AQ-27 were originally developed and evaluated within Western, English-speaking cultural contexts, such as the UK, USA and Australia, and based on theories that reflect Western assumptions and values.¹⁷ Cultural adaptation is as important as linguistic adaptation, but is arguably a somewhat more elusive ambition. It is likely that this will inform the way in which mental health is conceptualised and represented, and may potentially mean that meaningful efforts to develop.⁶² A report by the Lancet Commission¹¹ highlighted concerns that within the field of global mental health, Western, biomedical models of mental health are being extrapolated to define health, illness and treatment across diverse cultural contexts where a variety of different perspectives may be held.⁶³ An alternative approach could be to develop culturally specific stigma measures; Yang et al.¹⁷ propose a ‘what matters most’ framework to guide the development of culture-specific measures, which focuses on attempting to understand how stigma threatens the activities that define personhood within the local cultural context. This approach may be better able to capture culture-specific stigma dynamics.

Declarations of interest

None.

Ethical considerations

All analyses were based on previously published studies, therefore no ethical approval or patient consent is required.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.ejpsy.2025.100290](https://doi.org/10.1016/j.ejpsy.2025.100290).

References

- Thornicroft G, Brohan E, Rose D, Sartorius N, Leese M. Global pattern of experienced and anticipated discrimination against people with schizophrenia: a cross-sectional survey. *Lancet*. 2009;373:408–415. [https://doi.org/10.1016/S0140-6736\(08\)61817-6](https://doi.org/10.1016/S0140-6736(08)61817-6).
- Ciciurkaite G, Pescosolido BA. Mental Health Literacy and Public Stigma: examining the Link in 17 Countries. *Med Res Arch*. 2024;12(7). <https://doi.org/10.18103/mra.v12i7.5471>.
- Goffman E. *Stigma: Notes on the Management of Spoiled Identity*. New York: Simon and Schuster; 1963.
- Link BG, Phelan JC. Conceptualizing stigma. *Annu Rev Sociol*. 2001;27:363–385. <http://www.jstor.org/stable/2678626>.
- Kleinman A, Hall-Clifford R. Stigma: a social, cultural and moral process. *J Epidemiol Community Health*. 2009;63(6):418–419. <https://doi.org/10.1136/jech.2008.084277>.
- Yang LH, Kleinman A, Link BG, Phelan JC, Lee S, Good B. Culture and stigma: adding moral experience to stigma theory. *Soc Sci Med*. 2007;64(7):1524–1535. <https://doi.org/10.1016/j.socscimed.2006.11.013>.
- World Health Organisation. Comprehensive mental health action plan 2013–2030. <https://www.who.int/publications/i/item/9789240031029>; 2021 [accessed 16 August 2024].
- Thornicroft G, Sunkel C, Alikhon Aliev A, et al. The Lancet Commission on ending stigma and discrimination in mental health. *Lancet*. 2022;400(10361):1438–1480. [https://doi.org/10.1016/S0140-6736\(22\)01470-2](https://doi.org/10.1016/S0140-6736(22)01470-2).
- Pescosolido BA. Stigma as a mental health policy controversy: positions, options, and strategies for change. In: Goldman H, Frank R, Morrissey J, eds. *The Palgrave Handbook of American Mental Health Policy*. Cham: Palgrave Macmillan; 2020. https://doi.org/10.1007/978-3-030-11908-9_19.
- Krendl AC, Pescosolido BA. Countries and cultural differences in the stigma of mental illness: the East–West divide. *J Cross-Cultural Psychol*. 2020;51(2):149–167. <https://doi.org/10.1177/0022022119901297>.
- Patel V, Saxena S, Lund C, et al. The lancet commission on global mental health and sustainable development. *Lancet*. 2018;392:1553–1598. [https://doi.org/10.1016/S0140-6736\(18\)31612-X](https://doi.org/10.1016/S0140-6736(18)31612-X).
- Thornicroft G, Mehta N, Clement S, et al. Evidence for effective interventions to reduce mental-health-related stigma and discrimination. *Lancet*. 2016;387:1123–1132. [https://doi.org/10.1016/S0140-6736\(15\)00298-6](https://doi.org/10.1016/S0140-6736(15)00298-6).
- Bos AER, Pryor JB, Reeder GD, Stutterheim SE. Stigma: advances in theory and research. *Basic Appl Soc Psych*. 2013;35(1):1–9. <https://doi.org/10.1080/01973533.2012.746147>.
- Fox AB, Earnshaw VA, Taverna EC, Vogt D. Conceptualizing and measuring mental illness stigma: the mental illness stigma framework and critical review of measures. *Stigma Health*. 2018;3(4):348–376. <https://doi.org/10.1037/sah0000104>.
- Becker G, Arnold R. Stigma as a social and cultural construct. In: Ainlay SC, Becker G, Coleman LM, eds. *The Dilemma of difference: A multidisciplinary View of Stigma*. Springer; 1986:39–57.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417–1432. [https://doi.org/10.1016/0895-4356\(93\)90142-n](https://doi.org/10.1016/0895-4356(93)90142-n).
- Yang LH, Thornicroft G, Alvarado R, Vega E, Link BG. Recent advances in cross-cultural measurement in psychiatric epidemiology: utilizing ‘what matters most’ to identify culture-specific aspects of stigma. *Int J Epidemiol*. 2014;43(2):494–510. <https://doi.org/10.1093/ije/dyu039>.
- Corrigan P, Markowitz FE, Watson A, Rowan D, Kubiak MA. An attribution model of public discrimination towards persons with mental illness. *J Health Soc Behav*. 2003;44(2):162–179. <https://doi.org/10.2307/1519806>.
- Corrigan P. W. A toolkit for evaluating programs meant to erase the stigma of mental illness. [Internet]. 2015 May 9. Available from: <https://medbox.org/pdf/5e148832db60a2044c2d549e>.
- Corrigan PW, Powell KJ, Michaels PJ. Brief battery for measurement of stigmatizing versus affirming attitudes about mental illness. *Psychiatry Res*. 2014;215(2):466–470. <https://doi.org/10.1016/j.psychres.2013.12.006>.
- Hazell CM, Berry C, Bogen-Johnston L, Banerjee M. Creating a hierarchy of mental health stigma: testing the effect of psychiatric diagnosis on stigma. *BJPsych Open*. 2022;8(5):174. <https://doi.org/10.1192/bjo.2022.578>.
- Robinson P, Turk D, Jilka S, Cella M. Measuring attitudes towards mental health using social media: investigating stigma and trivialisation. *Soc Psychiatry Psychiatr Epidemiol*. 2019;54(1):51–58. <https://doi.org/10.1007/s00127-018-1571-5>.
- Corrigan PW. Mental health stigma as social attribution: implications for research methods and attitude change: science and practice. *Clin Psychol*. 2000;7(1):48–67. <https://doi.org/10.1093/clipsy.7.1.48>.
- Valdez D, Montenegro MS, Crawford BL, Turner RC, Lo WJ, Jozkowski KN. Translation frameworks and questionnaire design approaches as a component of health research and practice: a discussion and taxonomy of popular translation frameworks and questionnaire design approaches. *Soc Sci Med*. 2021;278, 113931. <https://doi.org/10.1016/j.socscimed.2021.113931>.
- 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;17:2701–2728. <https://doi.org/10.2147/JMDH.S419714>. Published 2024 May 31.
- Danielsen AK, Pommegaard HC, Burcharth J, Angenete E, Rosenberg J. Translation of questionnaires measuring health related quality of life is not standardized: a literature based research study. *PLoS one*. 2015;10(5), e0127050. <https://doi.org/10.1371/journal.pone.0127050>.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>.
- Acquadro C, Conway K, Hareendran A, Aaronson N. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health*. 2008;11(3):509–521. <https://doi.org/10.1111/j.1524-4733.2007.00292.x>.
- Eberhard DM, Simons GF, Fennig CD, eds. *Ethnologue: Languages of the World*. 26th ed. SIL International; 2023.

30. Kenny D. Human and machine translation. In: Kenny D, ed. *Machine Translation for everyone: Empowering users in the Age of Artificial Intelligence*. Berlin: Language Science Press; 2022:23–49. <https://doi.org/10.5281/zenodo.6759976>.
31. Liu W. The changing role of non-English papers in scholarly communication: evidence from Web of Science's three journal citation indexes. *Learn Publ*. 2016;30(2):115–123. <https://doi.org/10.1002/leap.1089>.
32. Scientific publishing has a language problem. *Na Hum Beh*. 2023;7(7):1019–1020. <https://doi.org/10.1038/s41562-023-01679-6>.
33. Mokkink L.B., Prinsen C.A.C., Patrick D.L., et al. COSMIN study design checklist for patient-reported outcome measurement instruments; 2019. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf.
34. Schellingerhout JM, Heymans MW, Verhagen AP, de Vet HC, Koes BW, Terwee CB. Measurement properties of translated versions of neck-specific questionnaires: a systematic review. *BMC Med Res Methodol*. 2011;11(87). <https://doi.org/10.1186/1471-2288-11-87>.
35. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>.
36. Popay J, Roberts H, Sowden A, et al. *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A product from the ESRC Methods Programme*. Lancaster University; 2006, 2006. Available from <https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersi on1-April2006.pdf>.
37. Meyers M, Geldmacher J, Mattausch S, et al. Stigmatisierung psychischer Erkrankung unter Schülern [Stigma of mental illness among students]. *Nervenarzt*. 2017;88(11):1266–1272. <https://doi.org/10.1007/s00115-016-0189-7>.
38. Pereira AdA, Santos SME, de Faria RMD. Versão brasileira do Attribution Questionnaire—Adaptação transcultural e validação de propriedades psicométricas [Brazilian version of the Attribution Questionnaire—Cross cultural adaptation and validation of psychometric properties]. *J Bras Psiquiatr*. 2016;65(4):314–321. <https://doi.org/10.1590/0047-2085000000139>.
39. Rodríguez-Meirinhos A, Antolín-Suárez L. Estigma social hacia la enfermedad mental: factores relacionados y propiedades psicométricas del Cuestionario de Atribución-revisado [Social stigma towards mental illness: related factors and psychometric properties of the revised-Attribution Questionnaire]. *Univ Psychol*. 2020;19:1–13. <https://doi.org/10.11144/Javeriana.upsy19.esem>.
40. Saavedra J, Murvartian L. Estigma público en salud mental en la universidad [Mental health public stigma at the university]. *Univ Psychol*. 2021;20:1–15. <https://doi.org/10.11144/Javeriana.upsy20.epsm>.
41. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annu Int Med*. 2009;151(4):65–94. <https://doi.org/10.1016/j.jclinepi.2009.06.006>.
42. Ayuso LS, Ruiz-Hontangas A, Cervantes JGG, et al. The promotion of mental health and prevention of first-episode psychosis: a pilot and feasibility non-randomised clinical trial. *Int J Environ Res Public Health*. 2023;20(22):7087. <https://doi.org/10.3390/ijerph20227087>.
43. González-Sanguino C, Muñoz M, Castellanos MA, Pérez-Santos E, Orihuela-Villameriel T. Study of the relationship between implicit and explicit stigmas associated with mental illness. *Psychiatry Res*. 2019;272:663–668. <https://doi.org/10.1016/j.psychres.2018.12.172>.
44. Muñoz M, Guillén AI, Pérez-Santos E, Corrigan PW. A structural equation modeling study of the Spanish mental illness stigma Attribution Questionnaire (AQ-27-E). *Am J Orthopsychiatr*. 2015;85(3):243–249. <https://doi.org/10.1037/ort0000059>.
45. Chamorro Coneo A, Aristizabal E, Hoyos de los Rios O, Aguilar D. Danger appraisal and pathogen-avoidance mechanisms in stigma towards severe mental illness: the mediating role of affective responses. *BMC Psychiatry*. 2022;22(330). <https://doi.org/10.1186/s12888-022-03951-x>.
46. Crespo M, Pérez-Santos E, Muñoz M, Guillén AI. Descriptive study of stigma associated with severe and persistent mental illness among the general population of Madrid (Spain). *Community Ment Health J*. 2008;44(6):393–403. <https://doi.org/10.1007/s10597-008-9142-y>.
47. Chiu YH, Kao MY, Goh KK, Lu CY, Lu ML. Renaming schizophrenia and stigma reduction: a cross-sectional study of nursing students in Taiwan. *Int J Environ Res Public Health*. 2022;19(6):3563. <https://doi.org/10.3390/ijerph19063563>.
48. Ho AHY, Fong TCT, Potash JS, Ho VFL, Chen EYH, Ho RTH. Deconstructing patterns of stigma toward people living with mental illness. *Soc Work Res*. 2018;42(4):302–312. <https://doi.org/10.1093/swr/svy022>.
49. Pingani L, Forghieri M, Ferrari S, et al. Stigma and discrimination toward mental illness: translation and validation of the Italian version of the Attribution Questionnaire-27 (AQ-27-I). *Soc Psychiatry Psychiatr Epidemiol*. 2012;47(6):993–999. <https://doi.org/10.1007/s00127-011-0407-3>.
50. Saguem BN, Gharmoul M, Braham A, Nasr SB, Qin S, Corrigan P. Stigma toward individuals with mental illness: validation of the Arabic version of the Attribution Questionnaire in a university student population. *J Public Ment Health*. 2021;20(3):201–209. <https://doi.org/10.1108/JPMH-10-2020-0135>.
51. Romem P, Anson O, Kanat-Maymon Y, Moisa R. Reshaping students' attitudes toward individuals with mental illness through a clinical nursing clerkship. *J Nurs Educ*. 2008;47(9):396–402. <https://doi.org/10.3928/01484834-20080901-01>.
52. Akyurek G, Efe A, Kayihan H. Stigma and discrimination towards mental illness: translation and validation of the Turkish version of the Attribution Questionnaire-27 (AQ-27-T). *Community Ment Health J*. 2019;55(8):1369–1376. <https://doi.org/10.1007/s10597-019-00438-0>.
53. Baminiwatta A, Alahakoon H, Herath NC, Kodithuwakku KM, Nanayakkara T. Trait mindfulness, compassion, and stigma towards patients with mental illness: a study among nurses in Sri Lanka. *Mindfulness (N Y)*. 2023;14(4):979–991. <https://doi.org/10.1007/s12671-023-02108-5>.
54. Giasuddin NA, Levav I, Gal G. Mental health stigma and attitudes to psychiatry among Bangladeshi medical students. *Int J Soc Psychiatry*. 2015;61(2):137–147. <https://doi.org/10.1177/0020764014537237>.
55. Ihalainen-Tamlander N, Vähäniemi A, Löytyniemi E, Suominen T, Välimäki M. Stigmatizing attitudes in nurses towards people with mental illness: a cross-sectional study in primary settings in Finland. *J Psychiatr Ment Health Nurs*. 2016;23(6–7):427–437. <https://doi.org/10.1111/jpm.12319>.
56. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>.
57. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>.
58. Dancy CP, Reidy J. *Statistics Without Maths For Psychology*. 7th ed. Harlow: Pearson; 2017.
59. Tavakol M, Wetzel A. Factor analysis: a means for theory and instrument development in support of construct validity. *Int J Med Educ*. 2020;11:245–247. <https://doi.org/10.5116/ijme.5f96.0f4a>.
60. Behr D. Assessing the use of back translation: the shortcomings of back translation as a quality testing method. *Int J Soc Res Methodol*. 2017;20(6):573–584. <https://doi.org/10.1080/13645579.2016.1252188>.
61. Nussbaumer-Streit B, Klerings I, Dobrescu AI, et al. Excluding non-English publications from evidence-syntheses did not change conclusions: a meta-epidemiological study. *J Clin Epidemiol*. 2020;118:42–54. <https://doi.org/10.1016/j.jclinepi.2019.10.011>.
62. Oliveira S, Baggs E. *Psychology's WEIRD Problems*. Cambridge University Press; 2023. <https://doi.org/10.1017/9781009303538>.
63. Johnstone L, Boyle M, Cromby J., et al. The Power Threat Meaning Framework: towards the identification of patterns in emotional distress, unusual experiences and troubled or troubling behaviour, as an alternative to functional psychiatric diagnosis. <https://www.bps.org.uk/guideline/power-threat-meaning-framework-full-version>; 2018 [accessed 16 August 2024].