Forty years (1981-2021) assessing stigma with the Community Attitudes to Mental Illness (CAMI) scale: systematic review of psychometric properties

Albert Bernadàs
Universitat Autònoma de Barcelona

Eduardo Doval
Universitat Autònoma de Barcelona: Universitat Autonoma de Barcelona

Natalia Angraita-Osorio
Hospital del Mar Medical Research Institute: Institut Hospital del Mar d'Investigacions Mediques

Juan P. Sanabria-Mazo
Parc Sanitari Sant Joan de Deu

Ariadna Colomer-Carbonell
Parc Sanitari Sant Joan de Deu

Sara Evans-Lacko
LSE: The London School of Economics and Political Science

Graham Thornicroft
King's College London Institute of Psychiatry Psychology and Neuroscience

Juan Vicente Luciano (juanvicente.luciano@uab.cat)
Universitat Autònoma de Barcelona: Universitat Autonoma de Barcelona

Maria Rubio-Valera
Parc Sanitari Sant Joan de Deu

Research Article

Keywords: Stigma, mental health, systematic review, Community Attitudes toward the Mentally Ill.

Posted Date: September 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1715477/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at Systematic Reviews on April 14th, 2023. See the published version at https://doi.org/10.1186/s13643-023-02230-4.
Abstract

Background. The Community Attitudes to Mental Illness (CAMI) scale measures social stigma towards people with mental illness. Although it has been used worldwide, the psychometric properties of the CAMI have not been systematically reviewed. The main aim of the present work was to systematically review the psychometric properties (dimensionality, reliability, and construct validity) of the different versions of the CAMI 40 years after its publication.

Methods. A systematic search was conducted in PubMed, PsycINFO, Web of Science, and EMBASE from January 1981 to June 2021. A double review was performed for eligibility, data extraction, and quality assessment.

Results. A total of 13 studies enrolling 10,095 participants were included. The most frequently reported factor structure comprises 3 or 4 factors. The internal consistency seems adequate for the global scale ($a \geq 0.80$), except for CAMI-10 ($a = 0.69$). Overall, internal consistency of the subscales is not supported, with authoritarianism being the weakest factor ($a = 0.27$ to 0.68). The stability over time of the total scale has been assessed in the CAMI-BR and CAMI-10 ($r > 0.65$). Few studies have assessed the temporal stability of the CAMI subscales. Most of the correlations with potentially related measures are significant and in the expected direction.

Conclusions. The 3-4 factor structure is the most widely reported in the different versions of the CAMI. Although the reliability and construct validity are acceptable, further item refinement by international consensus seems warranted 40 years after the original publication.

PROSPERO registration number: CRD42018098956

Introduction

Stigma towards people with mental disorders is a sociocultural phenomenon [1, 2] that negatively affects quality of life, self-esteem, interpersonal relationships, health care seeking and provision, and workplace integration [3, 4]. In Europe, it is estimated that there are about 165 million people with mental disorders, and it is calculated that around 38% of people will experience a mental disorder in their lifetime [5]. A recent systematic review confirmed that mental disorders cause a substantial economic burden for societies, with developmental disorders, schizophrenia and intellectual disabilities obtaining the top median societal cost per patient [6].

Although there is not a universally accepted definition of stigma, it can be considered a multidimensional construct composed of negative elements of knowledge, attitudes, feelings, beliefs, and behaviours towards a group of people [7]. Stigma is a powerful barrier to social participation and professional help-seeking for people with mental disorders and there is a widespread social belief that such people are aggressive and uncontrollable [8–12]. In the last decade, various countries have implemented anti-discrimination campaigns to reduce stigma and improve the integration of people with mental illness into communities [13]. Examples of such programmes are Time to Change in England [14]; Obertament in Spain [15]; Schizophrenia has many faces in Austria16; Like minds like mine in New Zealand [17]; One of us in Denmark [18]; and Opening minds in Canada [19].

There is a long history of scales developed to assess attitudes towards mental illness. The Opinion about Mental Illness (OMI) [20] and the Custodial Mental Illness Ideology Scale (CMI) [21] were developed in the 1950s-1960s as the first scales to measure stigma. More recently, the Community Mental Health Ideology (CMHI) [22], the Community Attitudes to Mental Illness Scale (CAMI) [23], the Mental Health Knowledge Schedule (MAKS) [24], and the California Assessment of Stigma Change (CASC) [25] were designed. Since its publication in 1981, the CAMI [23] has been the gold standard measure for assessing stigma towards people with mental disorders. It has been translated into several languages (Spanish, Italian, Swedish, Portuguese, and Greek, among others) and used to measure stigma in a wide variety of samples (e.g., nurses, psychiatrists, and relatives of psychiatric patients).

The original version of the CAMI was partially derived from a brief, revised, and updated version of the OMI [20], and it was initially developed to predict the reactions of the general population to local services for people with severe mental disorders. This original version is composed of 40 items that are responded on a 5-point Likert scale, ranging from "strongly agree" to "strongly disagree". According to its developers, the CAMI contains four subscales: authoritarianism, benevolence, social restrictiveness, and community mental health ideology [23]. Each subscale contains 10 items (5 positively formulated plus 5 negatively formulated) on the opinions of treating and caring for people with a serious mental disorder. Therefore, subscale scores can range from 10 to 50, with higher scores indicating less stigma towards people with mental disorders.

Given that the CAMI has been available for four decades, with hundreds of citations, the time is right for a systematic review of the psychometric properties of its different versions. As far as we know, there are no previous reviews summarizing available psychometric information on the CAMI. This systematic review bridges this gap by synthesizing and critically appraising the psychometric properties of this well-known stigma scale.

Method
Protocol and registration

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines (PRISMA) [26]. The protocol was registered with PROSPERO on July 25th, 2018 (registration number: CRD42018098956).

Search strategy

The search was conducted through four electronic databases: PubMed, PsycINFO, Web of Science, and EMBASE. The search strategy included terms related to psychometrics (psychometrics OR factor analysis OR reliability OR intra-class OR test-retest OR internal consistency OR validity OR dimensionality OR sensitivity to change OR responsiveness OR sensitivity OR specificity) and to the original scale name (attitude* AND toward* AND mental* AND Ill*) OR "CAMI", found using keywords in all fields and in the Medical Subject Headings (MeSH). The search string used in PubMed is shown in the Supplementary Table 1. The references of included studies were screened by reverse citation search to identify studies not detected in the electronic searches.

Eligibility criteria

The search in the databases incorporated studies published in peer-reviewed journals from 1981 (when the original version of Taylor and Dear [23] was published) to June 2021. This systematic review included all studies that provided evidence on the psychometric properties (factor structure, internal consistency, temporal stability, convergent validity, or other indicators of reliability and validity) of the different versions of the CAMI. To ensure the rigour of the included studies, literature reviews, books, doctoral dissertations, commentaries, conference abstracts, study protocols, case reports, qualitative studies, non-peer-reviewed manuscripts, were excluded. In addition, non-English, non-Spanish or non-Italian papers were also excluded. No restrictions were placed on the characteristics of the participants and the type of sampling used in the search.

Data management and study selection

In the first phase, duplicate articles were removed using Mendeley. In the second phase, two reviewers (A.B. and N.A.-O.) independently assessed the articles based on their title and abstract according to the eligibility criteria. In the third phase, the full text of those articles that met the second phase was reviewed to verify compliance with the eligibility criteria. In the fourth phase, discrepancies in study selection were resolved with the help of two additional external reviewers (J.V.L. and M.R.-V.). In the fifth phase, relevant data were extracted from the selected documents with a standardised data extraction form and the respective quality assessment was carried out for each document.

Data extraction

Data extraction from the selected articles was performed independently (A.B. and N.A.-O.), using a template containing the following sections: authors and year of publication, country, CAMI version, study design, sample type, sample size, and psychometric results on the CAMI. The authors of the study were contacted to obtain additional information on the psychometric properties of the scale, if necessary.

Quality assessment

The quality of the included studies was assessed using the criteria proposed by Terwee et al. [27] for health measures. Each of the criteria is scored 2 if the criteria are fulfilled, 1 if they are partially fulfilled, and 0 if no criteria are fulfilled. The total score can range from 0 to 14. The quality assessment was carried out by the first author (A.B.), with the supervision of two authors (J.V.L. and M.R.-V.). Specifically, the following psychometric properties were assessed.

Content validity indicates whether the construct of interest is sampled by the questionnaire items. The measurement objective of the questionnaire and the target population must be explicitly defined. Furthermore, for this criterion to be met, it is necessary to develop the questionnaire items in consensus with the general population and stigma experts [27].

Factor structure refers to the dimensionality of the scale [27]. A score of 2 was given if an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA) had been performed on different samples or if the CFA had been calculated considering a theoretical model. A score of 2 was only given if the factor analyses supported the structure promoted by the authors. A score of 1 was awarded if only the EFA had been carried out, and also if the EFA supported the factor structure. A score of 0 was awarded if factor analysis has not been conducted or if EFA or CFA does not support the proposed dimensionality.

Internal consistency is used to indicate the degree of reliability of a scale. For health scales, Cronbach's $\alpha$ should be between 0.70 and 0.95 [27]. For greater rigour in the findings, the Nunnally and Bernstein [28] criteria were used in this study, which present scores above 0.80 as acceptable.

Test–retest reliability is a measure used to validate the stability of the scale over time. A minimum of 1–2 weeks between the first and second administration is recommended to reduce possible recall bias. For acceptable temporal stability, the test-retest needs to be at least $r = 0.70$ [29]. The intraclass correlation coefficient (ICC) is the most recommended statistical index for continuous measures in the assessment of temporal stability.
Convergent and discriminant validity addresses whether scores on a questionnaire are significantly associated with potentially related measures. A theoretical underpinning is needed to verify the hypotheses of expected correlations between different scales. At least two of the correlations between two theoretically related constructs had to have a minimum of $r = 0.50$ [27].

Floor and ceiling effects is a measure to detect the number of participants achieving the highest or lowest possible scores. No more than 15% of the sample must obtain the maximum or minimum score on the questionnaire to meet these criteria [27].

Interpretability (i.e., how differences in scores on the CAMI can be interpreted, or the degree to which qualitative meaning can be obtained from quantitative scores). A known-groups validity approach is suggested with means and SDs of scores of relevant subgroups of subjects who are expected to differ in the CAMI [27].

Results

Selection and inclusion of studies

As displayed in Fig. 1, the initial database search yielded a total of 515 published articles. In addition, 5 articles were included by reverse citation and 5 by experts. After removal of duplicates, 493 titles and abstracts were reviewed, of which 22 were selected for full-text review. After this review, 9 articles were excluded, 4 articles because they were not related to the CAMI, 2 because they did not focus on the psychometric properties of the scale, 2 because they did not provide relevant information, and 1 because it was written in German. Finally, a total of 13 studies were included in this systematic review.

Characteristics of included studies

The 13 included studies were conducted in 10 different countries: Spain ($n = 2$), UK ($n = 2$), Italy ($n = 1$), France ($n = 1$), Sweden ($n = 1$), Ireland ($n = 1$), Canada ($n = 1$), China ($n = 1$), Chile ($n = 1$), and Argentina ($n = 1$). Participants in most studies were healthy individuals ($n = 12, 92\%$) recruited through non-probability sampling ($n = 10, 77\%$).

The original CAMI-40 ($n = 3$) [23, 30, 31] and the CAMI-20 ($n = 3$) [31–33] were the most psychometrically analysed versions, followed by CAMI-W ($n = 2$) [31, 34], CAMI-26 ($n = 2$) [15, 35], CAMI-BR ($n = 1$) [36], CAMI-31 ($n = 1$) [37], CAMI-10 ($n = 1$) [38], CAMI-24 ($n = 1$) [39], CAMI-22 ($n = 1$) [40], CAMI-1 [31], CAMI-2 [31], and CAMI-W ($n = 1$) [31]. Table 1 provides a detailed description of the included studies.
<table>
<thead>
<tr>
<th>Author (year); country</th>
<th>CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M + SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taylor &amp; Dear (1981) [23]; Canada</td>
<td>CAMI-40 (40 items)</td>
<td>Observational</td>
<td>General population from Toronto with separate samples from areas with and without mental health facilities (non-probabilistic sample)</td>
<td>1090</td>
<td>NA</td>
<td>Less considerate attitudes toward the mentally ill: elder people, divorced people, users with children (not including the ones with children older than 18 years old), regular church attenders (depending on the religious determination). The Pentecostal and Greek Orthodox groups showed more authoritarian attitudes and the Pentecostal and Greek Orthodox had the least benevolent attitudes. More considerate attitudes toward the mentally ill: female respondents, higher sociodemographic status, and higher education. The users that had used mental health facilities or are related with friends that had used those facilities. The Baptists and Salvation Army showed fewer authoritarian attitudes and the Baptists with the United Church, showed more benevolent attitudes.</td>
<td>- Factor structure: benevolence (10 items), CMHI (10 items), social restrictiveness (10 items), authoritarianism (10 items). - Global internal consistency: NA - Subscales internal consistency: - Benevolence: $\alpha = 0.76$ - CMHI: $\alpha = 0.88$ - Social restrictiveness: $\alpha = 0.80$ - Authoritarianism: $\alpha = 0.68$ - Global test-retest reliability: NA - Subscale test-retest reliability: NA - Validity of subscales: A strong degree of correspondence between the factor scales and the theoretically ones. It is important to emphasize that the authoritarianism and social restrictiveness scales equally correlated with the first factor, ad to a lesser extent, with the fourth factor.</td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country</th>
<th>CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brockington et al. (1993) [37]; UK</td>
<td>CAMI-31 (31 items)</td>
<td>Observational</td>
<td>General population from the Malvern and Bromsgrove area (non-probabilistic sample).</td>
<td>2000</td>
<td>NA (NA); Bromsgrove (51%); NA (NA); Malvern (58%)</td>
<td>CMHI: (factor 1): Higher education and professional experience with people with mental health problems. Authoritarianism (factor 2): Lower scores are related with higher education. Benevolence (factor 3): The highest scores were found within the groups with higher education, knowledge of mental illness and an age between 55–64.</td>
<td>- Factor structure: Benevolence (NA), CMHI (NA), authoritarianism (NA) - Factor Scores: NA - Global internal consistency: NA - Subscale internal consistency: NA - Global test-retest reliability: NA - Subscale Test-retest reliability: NA - CMHI: NA - Authoritarianism: NA - Benevolence: NA - Validity: NA</td>
</tr>
<tr>
<td>Wolff et al. (1996) [34]; UK</td>
<td>CAMI-W (20 items)</td>
<td>Observational</td>
<td>Participants from London (non-probabilistic sample).</td>
<td>215</td>
<td>37 (13); 55%</td>
<td>Fear and exclusion were associated with having children in the domestic establishment (&lt; 18), higher age, and a lower occupational status. Social control was associated with a higher age, sociodemographic lower class, Asian, African or Caribbean descendant, lower educational level, with children in the household, lower occupational status, being divorced, widowed or separated and be aware of somebody with a mental health problem. Goodwill was associated with lower age, higher educational level, and people from ethnical origin.</td>
<td>- Factor structure: fear and exclusion (11 items), social control (6 items), and goodwill (3 items). - Global internal consistency: NA - Sub-scale internal consistency: NA - Global Test-retest reliability: NA - Sub-scale Test-retest reliability: N - Validity: - Predictive validity: Significant correlation between fear and exclusion, social control and goodwill, and behavioural intention items</td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country; CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Song et al. (2005) [40]; China CAMI-22</td>
<td>Observational</td>
<td>General population from Taiwan (stratified proportional sampling).</td>
<td>1203</td>
<td>42.3 (14.2); 51%</td>
<td>- Factor structure: benevolence (6 items). Rehabilitation in the community (6 items), non-authoritarianism (5 items), non-social restrictiveness (3 items), and normalization (2 items).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global internal consistency: NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Sub-scale internal consistency:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Benevolence: α = 0.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Rehabilitation in the community: α = 0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Non-authoritarianism: α = 0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Non-social restrictiveness: α = 0.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Normalization: α = 0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global test-retest reliability: NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale test-retest reliability: NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Validity: NA</td>
<td></td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country</th>
<th>CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
</table>
| Buizza et al. (2005) [39]; Italy | CAMI-24 (24 items) | Observational | General population from Brescia (probabilistic sample). | 280 | 48.3 (NA); 60% | - Physical distance and fear: They are associated with age > 61 years, being divorced/widowed/separated and never having participated in voluntary or social activities.  
- Social isolation: Associated with age > 41 years, higher education and being unemployed  
- Social responsibility and tolerance: there are no significative association between the social demographic variables and this factor. | - Factor structure: physical distance and fear (8 items), social distance and isolation (9 items), and social responsibility and tolerance (7 items).  
- Global internal consistency: NA  
- Subscale internal consistency: NA  
- Global test-retest reliability: NA  
- Sub-scale Test-retest reliability: N  
- Validity: NA |
| Högberg et al. (2008) [32]; Sweden | CAMI-20 (20 items) | Observational | Student nurses selected based on their easy availability (non-probabilistic sample) | 421 | 27.9 (7.5); NA | - Factor structure: open-minded and pro-integration (9 items), fear and avoidance (6 items), and community mental health ideology (5 items).  
- Global internal consistency: α = 0.90  
- Subscale internal consistency: NA  
- Global test-retest reliability: NA  
- Subscale test-retest reliability: NA | |

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country; CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morris et al. (2011) [31]; Ireland CAMI (40 items), re-specified CAMI 1 (NA), re-specified CAMI 2 (NA), Högb erg CAMI-20 (20 items), CAMI-W (20 items), Re-specified Wolff Scale (NA).</td>
<td>Observational</td>
<td>Nurses (non-probabilistic sample).</td>
<td>1242</td>
<td>40 (10); 66%</td>
<td>NA</td>
<td>- Factor structure:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- CAMI-40: authoritarianism (10 items), benevolence (10 items), social restrictiveness (10 items), and CMHI (10 items).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Re-specified CAMI 1: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Re-specified CAMI 2: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Högb erg CAMI-20: open-minded and pro-integration (9 items), fear and avoidance (6 items), and community mental health ideology (5 items).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Wolff CAMI-20: fear and exclusion (11 items), social control (6 items), and goodwill (3 items).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Re-specified Wolff scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global internal consistency: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale internal consistency: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global test-retest reliability: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale Test-retest reliability: NA</td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country, CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
</table>
| Abelha et al. (2015) [36], Brazil CAMI-BR (40 items) | Observational | General population from the Rio de Janeiro area (non-probabilistic sample) | 230 | 44.90 (3.31); 68.3% | NA | - Factor structure and number of items: social restrictiveness (10 items), benevolence (10 items), CMHI (10 items), and authoritarianism (10 items).  
- Global internal consistency: $\alpha = 0.84$.  
- Subscale internal consistency:  
  - Social restrictiveness: $\alpha = 0.76$  
  - Benevolence: $\alpha = 0.69$  
  - CMHI: $\alpha = 0.81$  
  - Authoritarianism: $\alpha = 0.35$  
- Global test-retest reliability ICC = 0.69  
- Subscale test-retest reliability:  
  - Social restrictiveness: ICC = 0.64  
  - Benevolence: ICC = 0.62  
  - CMHI: ICC = 0.54  
  - Authoritarianism: ICC = 0.37  
- Validity: NA |

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country; CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ochoa et al. (2016) [30]; Spain CAMI-40 (40 items)</td>
<td>Observational</td>
<td>Elementary school students (non-probabilistic sample).</td>
<td>150</td>
<td>15.23 (0.79); 51.33%</td>
<td>NA</td>
<td>- Global CAMI scores: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Factor structure: authoritarianism (10 items), benevolence (10 items), social restrictiveness (10 items), and CMHI (10 items).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global internal consistency:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- First assessment: $\alpha = 0.86$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Second assessment $\alpha = 0.90$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>* Data provided by Ochoa et al. [32]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale internal consistency:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Authoritarianism: $\alpha = 0.27$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Benevolence: $\alpha = 0.64$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Social restrictiveness: $\alpha = 0.67$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- CMHI: $\alpha = 0.81$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global test-retest reliability: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale test-retest reliability:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- CMHI: ICC = 0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Authoritarianism: ICC = 0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Benevolence: ICC = 0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Social restrictiveness: ICC = 0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Validity: NA</td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country, CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M + SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grandón et al. (2016) [38]; Chile CAMI-10 (10 items)</td>
<td>Observational</td>
<td>Two samples of the general population from Concepción (non-probabilistic simple)</td>
<td>749 (pooling 2 samples)</td>
<td>First sample: 39.13 (12.57); 55.6%</td>
<td>NA</td>
<td>- Factor structure: acceptance (5 items) and CMHI (5 items).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second sample: 41.91 (12.57); 55.7%</td>
<td></td>
<td>- Global internal consistency: $\alpha = 0.69$.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale internal consistency:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Acceptance: $\alpha = 0.61$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- CMHI: $\alpha = 0.66$.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global test-retest reliability: ICC = 0.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Sub-scale test-retest reliability:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- CMHI: ICC = 0.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Authoritarianism: ICC = 0.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Benevolence: ICC = 0.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Social restrictiveness: ICC = 0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Validity:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Factor 1 SDO – factor 1 CAMI (-0.25).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Factor 1 SDO – factor 2 CAMI (-0.30).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Factor 2 SDO – factor 1 CAMI (-0.25).</td>
<td></td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country; CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
</table>
| Rubio-Valera et al. (2016) [15]; Spain CAMI-26 (26 items) | Observational | General population (probabilistic sample). | 1019 | 48.04 (1.25); 51.24% | Factor 2 SDO – factor 2 CAMI (0.16). | Factor 2 SDO 
- Factor structure: authoritarianism (7 items), benevolence (6 items) and support for community mental health ideology (9 items).  
- Global internal consistency: NA  
- Subscale internal consistency:  
  - Authoritarianism: $\alpha = 0.54$  
  - Benevolence: $\alpha = 0.63$  
- Support for community mental health care: $\alpha = 0.72$  
- Global test-retest reliability: NA  
- Subscale test-retest reliability: NA  
- Validity: NA |

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country; CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M ± SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garcia et al. (2017) [35]; France CAMI-26 (26 items)</td>
<td>Observational</td>
<td>Undergraduate nursing students (non-probabilistic)</td>
<td>268</td>
<td>NA</td>
<td>Factor structure: benevolence (7 items), CMHI (5 items), authoritarianism (7 items), and restrictiveness (7 items).</td>
<td>- Global internal consistency: NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale internal consistency: NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global Test-retest reliability: ICC = 0.79.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale test-retest reliability:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- CMHI: ICC = 0.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Authoritarianism: ICC = 0.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Benevolence: ICC = 0.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Social restrictiveness: ICC = 0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Validity: CAMI is related to the housing control scale, RIBS, and MARKS.</td>
<td></td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out. Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.
<table>
<thead>
<tr>
<th>Author (year); country; CAMI version (items)</th>
<th>Study design</th>
<th>Target population</th>
<th>Sample size</th>
<th>Age (years M + SD); Gender (% female)</th>
<th>Results depending on sociodemographic variables</th>
<th>Results about CAMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tong et al (2020) [33]; China CAMI-20 (20 items)</td>
<td>Observational</td>
<td>Medical students (MS) and primary healthcare workers (PHW)</td>
<td>MS (n = 1,228)</td>
<td>MS 20.8 (1.6); 63.4%</td>
<td>NA</td>
<td>- Factor structure: benevolence (5 items), fear and exclusion (8 items), and support and tolerance (7 items).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PHW (n = 1,092)</td>
<td>PHW 36.3 (10.2); 69.6%</td>
<td></td>
<td>- Global internal consistency:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>α = 0.82 for the MS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>α = 0.85 for the PHW</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Subscale internal consistency (MS and PHW, respectively):</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Benevolence: α = 0.79; 0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fear and Exclusion: α = 0.79; 0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Support and tolerance: α = 0.74; 0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Global test-retest reliability (MS and PHW): ICC = 0.79; 0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Test-retest reliability of the items:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ICCs ranged from 0.29 to 0.61 in MS; 0.45 to 0.74 in the PHW</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Validity:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Convergent validity: Correlation with potentially related measures was not assessed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Discriminant validity: good discriminant validity between subgroups of interest (MS and PHW).</td>
</tr>
</tbody>
</table>

* In those psychometric studies where a factor analysis has been carried out, resulting, for the most part, in a number of different factors and items, it has been decided, in the CAMI section used, to indicate the version resulting from the factor analysis previously carried out.

Brockington et al. [37] – Not all the items are indicated due to their non-specification in the original article.

**Quality assessment**
As shown in Table 2, the overall methodological quality of the included studies was low. On a scale of 0 to 14 points, 7 studies (54%) scored 5 or less [31, 32, 34, 35, 37, 39, 40] and 5 (38%) scored 6–7 [23, 15, 30, 36, 38]. Only one recent study [33] scored higher than 10 on the quality assessment, which indicates that in general the psychometric properties of the CAMI have not been adequately assessed.

Table 2
Quality assessment of the included studies using the criteria proposed by Terwee et al. [27].

<table>
<thead>
<tr>
<th>Articles</th>
<th>Content validity</th>
<th>Factor structure</th>
<th>Internal consistency</th>
<th>Test-retest reliability</th>
<th>Construct validity</th>
<th>Floor/ceiling effects</th>
<th>Interpretability</th>
<th>Final score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taylor &amp; Dear (1981) [23]</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Brockington et al. (1993) [37]</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Wolff et al. (1996) [34]</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Song et al. (2005) [40]</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Buizza et al. (2005) [39]</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Höberg et al. (2008) [32]</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Morris et al. (2011) [31]</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Abelha et al. (2015) [36]</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Ochoa et al. (2016) [30]</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Grandón et al. (2016) [38]</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Rubio et al. (2016) [15]</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>García et al. (2017) [35]</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Tong et al. (2020) [33]</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>11</td>
</tr>
</tbody>
</table>

The quality of the included studies was assessed using the criteria proposed by Terwee et al. [27] for health measures. Each of the criteria is scored 2 if the criteria are fulfilled, 1 if they are partially fulfilled, and 0 if no criteria are fulfilled. The total score can range from 0 to 14.

Content validity

The construct of interest in all included studies (n = 13) was the assessment of attitudes towards people suffering from mental disorders. The developers of the CAMI generated part of the scale by extracting items from previously published measures [23], whereas the subsequent versions were adaptations with different length of the original scale.

Factor structure

The dimensionality of the CAMI was assessed in most of the included articles (n = 10). Two studies computed a CFA [32, 35], six opted for an EFA or principal component analysis [15, 23, 34, 37, 38, 40], and 2 computed EFA and CFA [35, 36]. The number of dimensions ranged between 2 and 5. Regarding item allocation, it was found that items loaded on different subscales depending on the version of the CAMI. Considering suggested “Rules of thumb” [41], in the original CAMI (40 items distributed in 4 factors) only 26 of the 40 items had a factor loading greater than 0.40, with a difference ≥ 0.15 in the factor loadings of each item in the different factors. In addition, some items load more strongly on other factors than on the original factor assignment reported by Taylor & Dear [23].

Internal consistency

The internal consistency was assessed in 7 (54%) of the included studies. It was assessed for both the global CAMI (n = 5) [23, 30, 32, 36, 38] and its subscales (n = 6) [15, 23, 30, 36, 38, 40]. All studies that assessed the Cronbach's α of the global scale, both original CAMI-40 and its versions
The most reported structure

In this work, a total of 13 papers met the inclusion criteria and provided data on several psychometric indices. Although the 3-factor model was the different CAMI versions used, might account for the heterogeneous findings in the psychometric data of this popular stigma measure.

The study assessing the subscales of the CAMI-22 [40] reported internal consistency values below the minimum recommended cut-off point on the CAMI subscales.

On the other hand, there were studies that reported a Cronbach’s α value below the recommended cut-off only in some specific subscales of the original CAMI-40 [23, 30] and the CAMI-BR [36]. The two studies [23, 30] that evaluated the original CAMI indicated values below the cut-off on the dimensions of benevolence (α = 0.64 [23]; α = 0.76 [30]); and authoritarianism (α = 0.27 [23]; α = 0.68 [30]), and social restrictiveness (α = 0.67 [30]). The study [36] focused on the subscales of the CAMI-BR reported very low alpha values on benevolence (α = 0.69), authoritarianism (α = 0.35). In general, authoritarianism is the least reliable subscale [23, 30, 36, 38, 40].

Regarding the original CAMI, the two studies [23, 30] that assessed its subscales reported an acceptable internal consistency on the CMHI subscale (α = 0.88; α = 0.81, respectively), while only one [23] of the two studies reported an adequate value in social restrictiveness (α = 0.80). The study [36] that assessed the reliability of the CAMI-BR subscales also reported an adequate value on the CMHI subscale (α = 0.81).

Test-retest reliability

The temporal stability was assessed in only three studies [30, 35, 36] by computing the ICC coefficient. One study that used the original CAMI [30] calculated temporal stability on all subscales, without reporting the global coefficient: benevolence (r = 0.85), authoritarianism (r = 0.81), social restrictiveness (r = 0.81), and CMHI (r = 0.88). The study assessing temporal stability on the CAMI-BR [36] found acceptable temporal stability for the overall scale (r = 0.69). However, the subscales of benevolence (r = 0.62), authoritarianism (r = 0.37), social restrictiveness (r = 0.64), and CMHI (r = 0.54) did not exhibit good temporal stability. The temporal stability of the CAMI-10 [38] was acceptable only for the overall scale (r = 0.79), but not in the case of the subscales: benevolence (r = 0.39), authoritarianism (r = 0.57), social restrictiveness (r = 0.62), and CMHI (r = 0.63).

Construct validity

Two studies found positive statistically significant correlations of low magnitude (< 0.50) between the CAMI and potentially related instruments such as the MAKs and the RIBS. One study [38] identified correlations between the factors of the CAMI-10 (i.e., CMHI, authoritarianism, benevolence, and social restrictiveness) and the Social Dominance Orientation (SDO), with values ranging from ~ 0.31 to 0.16. Other study [35] found negative correlations between the CAMI-26 and instruments such as the RIBS (-0.44), the MAKs (-0.30) and the RIHCS (-0.44).

Floor and ceiling effects

None of the included studies reported information on ceiling and/or floor effects.

Interpretability

The interpretability was analysed from a known-groups validity approach in several studies (n = 10) [15, 23, 30, 33–37, 39, 40]. For instance, older participants with low employment status and low social class have higher CAMI scores compared to younger participants and those with high employment and social status [31, 37]. Men also reported higher scores than women [15]. As expected, those participants who had undertaken volunteering or social activities scored lower than those who had not [39].

Discussion

Principal findings and interpretation

The results of this systematic review can be summarized as follows. The CAMI has been used in a wide variety of settings and in diverse samples from many different countries (i.e., Spain, UK, Italy, France, Sweden, Canada, China, Chile, and Argentina). The target population of included studies varied from students [30], primary healthcare workers [31, 33], to the general population [15, 34, 35, 37, 38, 39, 40]. These aspects, besides the different CAMI versions used, might account for the heterogeneous findings in the psychometric data of this popular stigma measure.

In this work, a total of 13 papers met the inclusion criteria and provided data on several psychometric indices. Although the 3-factor model was the most reported structure [15, 32, 34, 37, 39] followed by a 4-factor model [23, 35, 36], the items considerably varied in their expected allocation among studies. Only 3 out of 13 studies presented the same number of items [23, 30, 31].
There were two studies using the 40-item version that differed from the original dimensions proposed for the CAMI. One study [15] supported a 3-factor structure, whereas the other supported a 4-factor structure [35]. The studies with more dramatic changes in item allocation were those using the CAMI-24 [39], CAMI-10 [38], and CAMI-26 [15]. The estimation method of maximum likelihood can only be used in CFA when multivariate normality is met, but in the CAMI-24 study [39] this assumption was violated. Moreover, the different approaches to analyse dimensionality (principal component analysis, EFA, and CFA) and the diverse methods used to estimate the factor models might explain the heterogeneous findings in dimensionality.

Regarding the internal consistency of the CAMI, 7 out of 13 studies addressed this psychometric aspect. Unexpectedly, only 4 studies reported the α for the total scale, whereas 3 reported α values only for subscales. According to Nunnally & Bernstein criteria [28], those studies that assessed the internal consistency of the total scale (CAMI-BR [36], CAMI-40 [30], and CAMI-20 [32]) obtained adequate α values, except for one study with the CAMI-10 [38]. In contrast, the studies that computed Cronbach’s α of each sub-scale [15, 23, 30, 36, 38, 40], usually obtained values below the recommended cut-off of 0.80. More specifically, the authoritarianism subscale was the least reliable sub-scale [23, 30, 36, 38, 40].

The CAMI-40 was the only version in which temporal stability of its sub-scales was analysed, showing adequate stability over time. CAMI-BR [36] and CAMI-10 [38] did not present good stability in the subscales but had an acceptable test-retest on the global scale. In general, there is a lack of longitudinal studies, therefore this psychometric aspect has not been exhaustively addressed.

The pattern of correlations between the CAMI and other potentially related constructs was statistically significant and in the expected directions, partially supporting the construct validity. Following Terwee et al. criteria [27], some correlations were not of the expected magnitude.

Finally, the interpretability of the CAMI scale assessed by 9 studies highlights that young female participants who have undertaken volunteering activities and have high employment and / or social status, present less stigma towards people with mental disorders.

**Strengths and limitations**

As far as we know, this systematic review is the first to summarize the psychometric properties of the different versions of the CAMI since it was published 40 years ago. The CAMI has been administered to a wide variety of populations. We applied adapted COSMIN criteria to evaluate the quality of the CAMI measurement properties; and provided a comprehensive and qualitative synthesis of its current evidence. For transparency purposes, we registered the review protocol in PROSPERO, and we have carried out a comprehensive search strategy as well as a clear data extraction procedure. In contrast, the scope of our results might be considered as limited due to the exclusion of papers not written in English, Spanish, or Italian. The relatively few included studies in this systematic review underly the need of addressing with more emphasis some psychometric properties of this popular stigma instrument, such as the factorial invariance across gender or cultures.

**Conclusion**

The CAMI psychometric properties have been examined using classic test theory as a framework. This methodological approach does not allow an assessment of the quality of individual CAMI items and response options across different levels of stigma. Methods based on item response theory would provide detailed information about the functioning of each CAMI item and would allow assessment of differential item functioning across gender or other sociodemographic variables. Although this scale has been employed in many samples from different cultures and with different languages, some aspects have been scarcely addressed. The time needed for completion, difficulties in understanding the items, or the scale’s acceptability have not been explored. As stated before, measurement invariance has not been assessed in its 40 years of history. Additionally, as only a few of the included studies had a longitudinal design, we cannot draw firm conclusions about the temporal stability of the CAMI. In our opinion, the next step should focus on item refinement to create a uniform set of items. This task would imply collaboration among an international panel of stigma experts. The resulting candidate scale should be evaluated by using cognitive interviews and surveying different samples to reach a final version of the scale with adequate dimensionality, reliability and validity that would allow cross-cultural comparisons.

**Declarations**

**Declarations of Interest.** The authors declare no conflict of interests.

**Data Availability Statement.** All data included in this systematic review were extracted from published papers.

**Author contributions.** E.D., J.V.L., S.E.-L., G.T., and M.R.-V. designed the study. A.B., J.P.S.-M., and A.C.-C. drafted the manuscript. A.B. and N.A.-O. contributed to the eligibility, systematic data extraction, and quality assessment. A.B., E.D., N.A.-O., J.P.S.-M., A.C.-C., J.V.L., S.E.-L., G.T., and M.R.-V. critically revised the manuscript. All authors commented on, revised and approved the draft and final manuscript.

**Funding.** JPS-M has a PFIS predoctoral contract from the Institute of Health Carlos III (ISCIII; FI20/00034). AC-C has a FI predoctoral contract from AGAUR (FLB/00216). We are also grateful to the Biomedical Research Center Network for Epidemiology and Public Health (CIBERESP; ISCIII) for its logistic support. This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit
References


Figure 1

PRISMA flowchart from record identification to study inclusion.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PRISMA2020checklistSB.docx
- Reference.pdf
- SupplementaryTable1.docx