Adaptation and Psychometric Evaluation of the Indonesian Version of the Multigroup Ethnic Identity Measure–12 Among Minangkabau Adults

Ira Syofyanti[1], Cleoputri Yusainy*[1], Sumi Lestari[1]
Hansel Jastissio Nerus Samosir[2], Indri Oktavia[1]
[1]Universitas Brawijaya, Indonesia
[2]PschStatBuddy, Indonesia
cleo.yusainy@ub.ac.id*

Submitted:
2026-09-14

Revised:
2026-10-16

Published:
2026-11-06

Keywords:
Confirmatory Factor Analysis, Ethnic Identity, Measurement Invariance, Response Process Validity, Scale Adaptation; Content Validity

Copyright holder:
© Author/s (2026)

This article is under:

How to cite:
Syofyanti, I., Yusainy, C., Lestari, S., Samosir, H. J. N., & Oktavia, I. (2026). Adaptation and Psychometric Evaluation of the Indonesian Version of the Multigroup Ethnic Identity Measure–12 Among Minangkabau Adults. *Bulletin of Counseling and Psychotherapy*, *8*(2). https://doi.org/10.51214/002026082307000

Published by:
Kuras Institute

E-ISSN:
2656-1050

ABSTRACT: The use of ethnic identity instruments in different cultural contexts requires evidence that the adapted version has content appropriateness, internal structure, and measurement equivalence. This study aimed to adapt the Multigroup Ethnic Identity Measure-12 (MEIM-12) to Indonesian and evaluate the psychometric evidence of the scores among Minangkabau adults. The adaptation process included forward translation, synthesis, back-translation, expert judgment by five experts, content validity assessment, and cognitive interviews. The internal structure was tested using confirmatory factor analysis with a maximum likelihood robust estimator to compare unidimensional and two-factor models, while reliability was evaluated using McDonald's omega and convergent validity using average variance extracted (AVE). Measurement invariance across sexes was tested using a multi-group confirmatory factor analysis. All items demonstrated adequate content validity (I-CVI ≥ .80; S-CVI/Ave = .95). The modified two-factor model showed the best fit (CFI = .944, TLI = .925, RMSEA = .083, SRMR = .056). All factor loadings were significant, and both dimensions demonstrated good reliability (ω = .806–.844). AVE met the criterion for the Affirmation and Belonging dimension (.512), whereas the Exploration dimension had a value of .456. Measurement invariance testing supported measurement equivalence up to the strict invariance level across the sexes. Overall, the Indonesian version of the MEIM-12 demonstrated preliminary evidence of validity and reliability, supporting its use as a two-dimensional measure of ethnic identity among Minangkabau adults. The instrument may also provide complementary information for culturally responsive counselling, particularly in exploring clients' experiences of ethnic identity and belonging.

INTRODUCTION

Indonesia is one of the world's most plural societies. The archipelago is inhabited by various ethnic groups with diverse regional languages, kinship systems, customs, and migration histories. This plurality is not merely a demographic backdrop but constitutes the social context in which individuals understand who they are, become part of a group, and how that group membership is

given meaning in everyday life. In this context, ethnic identity matters because it links personal experience with family background, social relations, and one's position within a multicultural society. Data from the 2020 Population Census also show that ethnic and regional language diversity remains a fundamental characteristic of the Indonesian population (Badan Pusat Statistik, 2024).

Ethnic identity refers to the self-concept that derives from knowledge of membership in an ethnic group, as well as the values and emotional significance attached to that membership. Within Phinney's framework, forming ethnic identity involves exploring the group's traditions, history, values, and practices, as well as commitment to, or a sense of belonging to, the ethnic group (Phinney, 1992; Phinney & Ong, 2007). Therefore, ethnic identity is not identical to the ethnic label a person holds. Two individuals of the same lineage may have different levels of exploration, attachment, and meaning-making of ethnic identity. This distinction matters in the Indonesian context because ethnic membership is often tied to language, region of origin, lineage, customs, and cross-regional life experiences.

Ethnic identity also does not develop within the static social dimension. Educational and occupational mobility, intergroup marriage, urbanization, and inter-island migration bring individuals into contact with environments that differ from their communities of origin. In a study of Indonesian students pursuing education abroad, (Nugraha, 2022) showed that the formation of identity cannot be adequately explained by geographical relocation alone; life experiences, social interactions, the adjustment process, and daily life at the destination constitute a more decisive context. Although the findings on students cannot be generalized to all adults, that study demonstrated that mobility can become a space for maintaining, negotiating, or reaffirming ethnic identity.

Likewise, globalization can broaden intergroup interaction, including the use of social media that enables the exchange of information regarding language, customs, group characteristics, and experiences of ethnic membership beyond geographical boundaries. Ariely (2019), showed that in societies with higher levels of globalization, lineage tends to be increasingly regarded as less of the sole basis for membership in a national group. This finding does not directly prove a weakening of individual ethnic identity but indicates that globalization can change how such group membership is understood. Therefore, globalization and social media are viewed more appropriately as new social spaces in which ethnic identity can be maintained, negotiated, displayed, or reinterpreted, rather than as the sole cause of weakening ethnic identity. These dynamics underscore the need for a measurement instrument that is not only translated linguistically but also functions adequately within the social context of the target ethnic group.

Minangkabau society offers a distinctive context for examining this issue. Minangkabau are known for their matrilineal kinship system, in which clan membership and the inheritance of ancestral property are passed down through the maternal line. In their social lives, the relationships among customs, Islam, the extended family, and the nagari form an important framework for membership and social responsibility (Handrianto, 2017). At the same time, the tradition of *merantau* (migration) leads many Minangkabau people to live outside West Sumatra and interact intensively with other ethnic groups (Kato, 1982; Parker, 2011). *Merantau* is not merely a change of residence, but also exposes individuals to a social environment that may differ from their community of origin. Thus, although Minangkabau membership based on the maternal line is relatively clear, the experience of exploring, maintaining, and expressing ethnic identity may vary by generation, place of residence, family environment, and "*merantau*" experience.

This variation is relevant to multicultural counseling practice. Understanding ethnic identity can help counselors situate a client's experience within the context of family, community, migration, belonging, and the meaning-making process of ethnic membership. However, ethnic identity scores must not be used to determine a person's authenticity as a Minangkabau or to make a clinical diagnosis. An adequate instrument remains necessary as a source of information in research. For

example, in a non-diagnostic assessment with interpretation that considers the social context of Minangkabau society, as well as validity evidence for a specific intended use.

One widely used instrument for assessing ethnic identity is the Multigroup Ethnic Identity Measure (MEIM), developed by Phinney (1992). The 12-item configuration used in this study assesses two interrelated dimensions: exploration and commitment, or affirmation, belonging, and commitment (Phinney & Ong, 2007). Although MEIM-12 is widely used, previous research has not always produced the same internal structure. Several studies support two correlated factors, while others have found a single-factor model or required reviewing items and residuals in certain populations. This variation indicates that the structure of a measurement instrument should not be assumed to be identical when applied to different languages, age groups, or ethnic groups.

Using a measurement instrument across cultures cannot be accomplished merely by translating words. Adaptation is necessary to examine construct, language, and response equivalence; ensure each item is relevant and comprehensible; and test the internal structure and consistency of scores in the target population. The International Test Commission guidelines emphasize that adaptation decisions must be supported by documentation of the translation process, expert review, assessment of respondents' comprehension, and psychometric evidence appropriate to the data characteristics (International Test Commission, 2017). In Indonesia, research on ethnic identity has developed across various populations, but evidence regarding the adaptation and psychometric evaluation of the MEIM-12, which specifically considers the social characteristics of Minangkabau society, remains limited. This limitation matters because the matrilineal system and the "*merantau*" experience may shape how ethnic membership, exploration, and commitment are understood.

This study aimed to (1) evaluate evidence of content validity and response process validity of the adapted version of the MEIM-12 among the Minangkabau society; (2) evaluate evidence of the internal structure validity and reliability of the MEIM-12; and (3) test the measurement invariance of the MEIM-12 across sexes.

METHODS

Design

This study used a psychometric design to adapt and evaluate the MEIM-12 among the Minangkabau society. The adaptation process was carried out following the stages of Beaton et al. (2000). These stages included initial translation by independent translators, synthesis of the translations, back-translation, expert checking and assessment of conceptual and meaning equivalence, and a final stage that tested the instrument in the population.

In the initial translation stage, two independent translation agencies translated the original scale. One agency was not given information about the psychological construct being measured, whereas the other assigned a translator with a psychological background who understood the construct. Using two agencies with different levels of understanding of the construct aimed to obtain a translation that addressed not only linguistic equivalence but also the appropriateness of each item's psychological meaning.

Subsequently, experts synthesized the results of these two translations to determine the word equivalents and sentence formulations most appropriate to the construct being measured and the scale's intended context. The synthesized results then underwent a back-translation stage into the scale's original language. This stage evaluated the consistency of concepts and meanings between the adapted version and the original scale.

In the fourth stage, all documents comprising the original version, initial translation, synthesis, and back-translation were provided to the expert judgment panel for examination and assessment of the equivalence of concepts, meanings, and language. The expert judgment involved five experts with complementary backgrounds relevant to the cross-cultural adaptation of MEIM-12. Three

experts had backgrounds in psychology, one in English and public communication, and one in academic English. All experts were of Minangkabau background and had knowledge and understanding of Minangkabau customs and culture. Their diverse expertise supported the evaluation of the conceptual, linguistic, and cultural appropriateness of the adapted items. This assessment ensured that the adapted version had conceptual and semantic equivalence with the original scale, while also fitting the cultural context and characteristics of the target population. The entire adaptation process was documented to support the interpretation and use of the MEIM-12 scores among the Minangkabau society.

Participants

The target population was Minangkabau society, including both men and women aged at least 18 years. The criterion for Minangkabau membership was established based on the matrilineal line, namely having a Minangkabau mother or both parents being Minangkabau. Participants were recruited using convenience and snowball sampling through an online questionnaire. We collected 525 responses in the initial stage, and 464 were available in the clean dataset. A total of 46 respondents were eliminated because they did not have a Minangkabau ethnic background, 8 respondents were eliminated because they were below the specified age criterion, 2 respondents were eliminated because they were unable to read/understand Indonesian, and 5 respondents were eliminated because they were indicated to have engaged in blocking when answering the questionnaire. The age range of respondents was 18–74 years (M = 29.00, SD = 12.15), with respondents predominantly female (58.85%), with a background of both Minangkabau mother and father among 413 respondents (89.01%), and only a Minangkabau mother among 51 respondents (10.99%).

Instrument

The MEIM-12 used in this study was based on the version of Roberts et al. (1999), which was developed from the original MEIM of Phinney (1992). This self-report scale measures ethnic identity in two interrelated dimensions. The exploration dimension consists of five items that assess an individual's efforts to seek information, learn about the group's background, and become involved in activities related to their ethnic group. The commitment dimension, which includes affirmation, belonging, and commitment, consists of seven items that assess clarity of membership, acceptance, sense of belonging, attachment, and commitment to the ethnic group.

Each item was rated using a four-category Likert scale: 1 (strongly disagree), 2 (disagree), 3 (agree), and 4 (strongly agree). Item composition, score computation, and final score interpretation followed the MEIM-12 factor structure obtained from psychometric testing among the Minangkabau society.

Cross-Cultural Adaptation Procedure

The MEIM-12 adaptation process involved initial translation, synthesis of translations, back-translation, meaning alignment, and assessment by five experts. We evaluated each item’s relevance using the item-level content validity index (I-CVI), modified kappa, and scale-level content validity index/average (S-CVI/Ave). The assessment results showed an S-CVI/Ave of .95, and all items obtained an I-CVI of at least .80. Every expert input, as well as the decision to retain or refine the wording of an item, was documented as part of the adaptation process.

We conducted cognitive interviews online from 8 August–14, 2026, with seven Minangkabau adults varying in age, sex, and place of residence. The interviews used a verbal probing technique. Participants were asked to restate the meaning of each item in their own words, describe the

experiences or considerations they used when answering, identify unclear words or phrases, and explain how they selected their response category.

The interview results showed that all items were generally understood in accordance with the intended construct of ethnic identity, and no substantive problems were found. Participants selected responses based on personal experience, knowledge acquired from family, interaction with fellow Minangkabau people, family habits, and "*merantau*" experience. One editorial input was found for item 3, related to the repetition of the word "*saya*" (I/me). The wording of that item was simplified without changing the meaning being measured, whereas the other items were retained. Participants also distinguished the four response categories fairly clearly. Several participants noted that older respondents, or those less accustomed to using online forms, might need more time to read, reflect, and determine their responses.

Data Collection and Ethical Considerations

Data collection was conducted online through Google Forms from 21 July to 10 August 2026. The instrument was distributed online by disseminating study information on social media and Minangkabau community networks. Before completing the questionnaire, participants received information about the researchers' identities, the study's aims and procedures, the participant criteria, the voluntary nature of participation, and data confidentiality. Participants who met the criteria and were willing to participate completed the questionnaire. Participants could discontinue completion without penalties. The data were used for research purposes and scientific reporting in an aggregate form.

Data Analysis

This study used CFA to evaluate construct validity, Average Variance Extracted (AVE) to evaluate convergent validity, and McDonald's omega to evaluate internal reliability. The analysis was carried out using the RStudio software developed by Posit Team (2026). *The packages used were lavaan (Rosseel et al., 2025)*, *dplyr* (Wickham et al., 2026), *semtools* (Jorgensen et al., 2026), and psych (Revelle, 2026). The analysis began by examining data completeness, response patterns, duplication, category distribution, and floor and ceiling effects. CFA was estimated using the maximum likelihood robust (MLR) estimator to accommodate the non-normality of item responses (Curran et al., 1996). Model fit was reported using the scaled chi-square, comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA) together with its 90% confidence interval, and standardized root mean square residual (SRMR), with criteria of CFI/TLI ≥ .90 (acceptable) and ≥ .95 (good), RMSEA ≤ .08, and SRMR ≤ .08 (Hu & Bentler, 1999; MacCallum et al., 1996). When the model fit was not yet adequate, modifications were considered based on modification indices accompanied by a theoretical justification in the form of the closeness of meaning between items. We tested measurement invariance across sexes in stages using configural, metric, scalar, and strict models, comparing models using the scaled chi-square difference test and change-in-fit criteria (ΔCFI ≤ .005 and ΔRMSEA ≤ .010) (Chen, 2007).

RESULTS AND DISCUSSION

Results

Content Validity and Response Process

The adaptation process of the MEIM-12 was carried out in several stages: forward translation, synthesis, back-translation, expert judgment, cognitive interview, revision, and finalisation (Table 1). The assessment results from the five experts showed that all items had an I-CVI ≥ .80, with an S-CVI/Ave value of .95. These results indicate that all items had adequate content relevance. Subsequently, cognitive interviews were conducted to evaluate participants' comprehension of the item content and response categories. In general, participants were able to understand the items in

accordance with the intended measurement, and no substantive problems were found in the response process. Revisions were made to items requiring refinement without changing the meaning of the construct being measured.

Table 1. Adaptation trail and cognitive interview

Stage	Activity	Result
1	Forward translation	Instrument translated into Indonesian
2	Synthesis & back-translation	Translation aligned with the original version
3	Expert judgment	Assessed by 5 experts
4	Content validity	I-CVI ≥ .80 for all items; S-CVI/Ave = .95
5	Cognitive interview	Conducted to evaluate comprehension of items and response categories
6	Revision	Items requiring improvement were refined without changing the meaning of the construct
7	Finalization	The adapted version was used for data collection

Internal Structure and Reliability

The examination of descriptive statistics for the twelve MEIM-12 items (Table 2) showed that most items had relatively low skewness and kurtosis values, with skewness ranging from −0.34 to −0.81 and kurtosis from −0.81 to 0.29, consistent with a level of non-normality that is still tolerable when using a robust estimator (Curran et al., 1996). Three items (M5, M6, and M9) showed higher skewness values (−1.61 to −1.94) and kurtosis values (2.72 to 4.40), accompanied by the highest median (4) and means approaching the upper limit of the scale (3.56–3.70). M12 also showed a tendency toward high-score responses, with a skewness of −1.00 and kurtosis of 0.97. The patterns in M5, M6, and M9 indicate a ceiling effect, namely, a tendency for responses to be concentrated in the highest score category. This finding was subsequently considered when testing the factor structure, particularly when evaluating the residual covariances between item pairs.

Table 2. Item Descriptive Statistics

Item	Mean	SD	Median	Min	Max	Skewness	Kurtosis
M1	3.19	0.72	3	1	4	-0.51	-0.22
M2	2.86	0.92	3	1	4	-0.34	-0.81
M3	3.15	0.72	3	1	4	-0.58	0.20
M4	3.00	0.82	3	1	4	-0.48	-0.34
M5	3.70	0.55	4	1	4	-1.94	4.40
M6	3.56	0.66	4	1	4	-1.66	2.93
M7	3.38	0.64	3	1	4	-0.63	-0.12
M8	3.07	0.79	3	1	4	-0.51	-0.26
M9	3.69	0.52	4	1	4	-1.61	2.72
M10	3.23	0.74	3	1	4	-0.63	-0.18
M11	3.36	0.68	3	1	4	-0.81	0.29
M12	3.52	0.59	4	1	4	-1.00	0.97

The factor structure of MEIM-12 was then tested through a comparison of three CFA models (Table 3). The unidimensional model showed an inadequate fit ($\chi^2(54) = 445.529$, $p < .001$; CFI = .819; TLI = .778; RMSEA = .143, 90% CI [.131, .155]; SRMR = .085), with CFI, TLI, and RMSEA not yet meeting the recommended fit criteria (CFI/TLI ≥ .95, RMSEA ≤ .06, SRMR ≤ .08; Hu & Bentler, 1999). The Correlated Two-Factor Model showed improved fit ($\chi^2(53) = 313.750$, $p < .001$; CFI = .882; TLI = .853; RMSEA = .116, 90% CI [.104, .129]; SRMR = .079), but the fit indices were still below the

established criteria. Four residual covariances were added on the basis of the modification indices and the item content. Three of them link M5, M6, and M9 (M5–M6, M5–M9, and M6–M9), which all load on the Affirmation and Belonging factor.

In the Indonesian version, these items use almost the same phrase, "*bagian dari (kelompok) etnis Minangkabau*" (being part of the Minangkabau ethnic group), and each describes a positive feeling about that membership: being happy (M5), having a strong sense of being part of the group (M6), and being very proud (M9). They also had the highest means and the most negative skewness of the 12 items (M = 3.56–3.70; skewness = -1.61 to -1.94; Table 2), so responses to all three clustered at the top of the scale. Because the three items share wording, emotional tone, and response distribution, they can carry variance that the Affirmation and Belonging factor does not explain.

The fourth covariance was specified between M3 and M4, although these items load on different factors (Affirmation and Belonging and Exploration, respectively). Both ask respondents to consider what Minangkabau membership means for them personally. M3 asks whether they clearly understand what their ethnic background means to them, and M4 asks how much they think about the way that membership may affect their lives. This common self-referential content may explain why the two items remain associated after the latent factors are taken into account Indonesian version, these items share almost the same phrase, "*bagian dari (kelompok) etnis Minangkabau*" (being part of the Minangkabau ethnic group), and each expresses a positive emotional evaluation of that membership: feeling happy (M5), having a strong sense of being part of the group (M6), and feeling very proud (M9).

These items also had the highest means and the strongest negative skewness among the 12 items (M = 3.56–3.70; skewness = -1.61 to -1.94; Table 2), indicating a shared ceiling pattern. This overlap in wording, affective tone, and response distribution may produce shared variance that is not explained by the Affirmation and Belonging factor. The fourth covariance was specified between M3 and M4, even though the two items load on different factors (Affirmation and Belonging and Exploration, respectively). Both items ask respondents to reflect on the personal relevance of Minangkabau membership: M3 concerns understanding what one's ethnic background means for oneself, whereas M4 concerns thinking about how that membership may affect one's life. This shared self-referential content may account for their residual association beyond the two latent factors.

The Modified Two-Factor Model produced an improved fit ($\chi^2(49) = 175.058$, $p < .001$; CFI = .944; TLI = .925; RMSEA = .083, 90% CI [.070, .097]; SRMR = .056). The CFI and TLI values in that model indicated an acceptable fit (≥ .90), whereas the SRMR was below .08. However, RMSEA remained slightly above the conventional .08 criterion despite showing substantial improvement compared with the two previous models (MacCallum et al., 1996). The Modified Two-Factor Model was therefore retained as the final measurement model because it showed the best overall fit among the models tested.

Table 3. Model Comparison

Model	χ^2 (scaled)	df	p-value	CFI[a]	TLI[a]	RMSEA [CI 90%][a]	SRMR
Unidimensional	445.529	54	<.001	0.819	0.778	0.143 [0.131, 0.155]	0.085
Correlated Two-Factor	313.75	53	<.001	0.882	0.853	0.116 [0.104, 0.129]	0.079
Modified Two-Factor	175.058	49	<.001	0.944	0.925	0.083 [0.070, 0.097]	0.056

Note. [a] Estimated using a robust estimator (MLR).

Further evaluation of the parameters of the Modified Two-Factor Model (Table 4) showed that all items had significant factor loadings ($p < .001$), with loadings ranging on the Exploration

dimension from 0.594 (M2) to 0.759 (M8) and on the Affirmation and Belonging dimension from 0.634 (M5) to 0.796 (M7). McDonald's omega for both dimensions was also good (Exploration = 0.806; Affirmation and Belonging = 0.844), exceeding the recommended threshold (≥ .70; Cheung et al., 2023). The AVE value for the Affirmation and Belonging dimension (.512) met the threshold for convergent validity (≥ .50; Fornell & Larcker, 1981), whereas the AVE value for the Exploration dimension (.456) was slightly below that threshold. Given that the reliability of this dimension exceeded .70, and all factor loadings were significant and above .50, the evidence for convergent validity of the exploration dimension should be interpreted with some caution..

Thus, the adapted version of the MEIM-12 for the Minangkabau society demonstrated reasonably adequate preliminary evidence in terms of internal structural validity and reliability for use as an instrument for measuring ethnic identity in this population. Nevertheless, the validity and reliability evidence at the overall sample level does not guarantee that this instrument functions equivalently across the subgroups within it. Therefore, the next analytical step was to test the equivalence of the measurement structure (measurement invariance) between the male and female groups through multi-group confirmatory factor analysis (MGCFA) to ensure that MEIM-12 scores can be validly compared across sex groups before being used further in research or non-diagnostic assessment.

Measurement Invariance Across Sex

The results of the multi-group confirmatory factor analysis (MGCFA) showed that the configural model, which served as the baseline model without equality constraints across groups, had a good fit with the data ($\chi^2(98) = 232.33$, CFI = 0.941, TLI = 0.920, RMSEA = 0.085, SRMR = 0.057) (Table 5). The difference test from the configural to the metric model ($\Delta\chi^2 = 8.48$, $\Delta df = 10$, $p = .582$) was not significant, indicating that the addition of constraints on the equality of factor loadings across groups did not significantly worsen the model. Likewise, comparisons from the metric to the scalar model ($\Delta\chi^2 = 11.45$, $\Delta df = 10$, $p = .324$) and from the scalar to the strict model ($\Delta\chi^2 = 14.07$, $\Delta df = 12$, $p = .296$) were not significant, indicating that the successive addition of constraints on the equality of intercepts and error variances did not significantly worsen the model fit.

These results were also supported by changes in goodness of fit that were far below the thresholds recommended for the condition of an unbalanced sample, with each group size below 300 (ΔCFI ≤ .005, ΔRMSEA ≤ .010; Chen, 2007). The changes in CFI across all testing stages (ΔCFI = 0.001 to -0.003) and RMSEA (ΔRMSEA = -0.003 to -0.004) were far below these thresholds,

Table 4. Factor Loadings and Reliability

Item	Loading (λ)	SE	z	p-value	CR (ω)	AVE
Exploration						
M1	0.692	0.034	14.51	<.001	0.806	0.456
M2	0.594	0.042	13.02	<.001		
M4	0.703	0.037	15.35	<.001		
M8	0.759	0.032	18.95	<.001		
M10	0.641	0.032	14.78	<.001		
Affirmation and Belonging						
M3	0.674	0.034	14.2	<.001	0.844	0.512
M5	0.634	0.034	10.33	<.001		
M6	0.725	0.036	13.22	<.001		
M7	0.796	0.027	18.52	<.001		
M9	0.680	0.031	11.38	<.001		
M11	0.760	0.030	17.22	<.001		
M12	0.695	0.029	14.02	<.001		

confirming that measurement invariance was not disturbed by the addition of increasingly stringent constraints. With equivalence achieved at each level, it can be concluded that the MEIM is fully invariant up to the strict level between the male and female groups. This finding indicates that, within this sample, the MEIM-12 demonstrated an equivalent factor structure, factor-loading pattern, item intercepts, and residual variances across the two sex groups.

Discussion

Content Validity and Response Process of the MEIM-12

The results of this study provide preliminary evidence that the process of adapting the MEIM-12 into Indonesian preserved the content appropriateness and comprehension of the items among Minangkabau society. The experts' assessment showed that the MEIM-12 items had adequate relevance to the construct to be measured, whereas the cognitive interviews provided additional evidence regarding how respondents understood and responded to those items. These two sources of evidence complement each other because the validity of an instrument needs to be supported by various sources of evidence, including evidence based on the test content and response process. Thus, the results at the early stage of adaptation support the use of the Indonesian version for further psychometric evaluation in a sample of Minangkabau society.

Table 5. MG-CFA by Sex

Model	χ^2 (scaled)	df	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	p	ΔCFI	ΔRMSEA
Configural	232.33	98	0.941	0.92	0.085	0.057					
Metric	241.37	108	0.942	0.929	0.08	0.061	8.48	10	0.582	0.001	-0.004
Scalar	254.57	118	0.941	0.934	0.077	0.063	11.45	10	0.324	-0.001	-0.003
Strict	262.28	130	0.94	0.939	0.074	0.068	14.07	12	0.296	-0.001	-0.003

Note. Δ was calculated relative to the preceding model sequentially (configural → metric → scalar → strict).

The high level of agreement among the experts indicates that the content of the Indonesian version of the MEIM-12 was generally regarded as relevant for representing the construct of ethnic identity in the target population. The involvement of experts with complementary psychological, linguistic, communication, and cultural expertise may have contributed to the evaluation of both the conceptual relevance and linguistic and cultural appropriateness of the adapted items. In the cross-cultural adaptation of an instrument, the evaluation of content validity is important to ensure that items are not only translated linguistically but also continue to represent the intended construct in the new cultural context. Linguistic equivalence alone does not always guarantee conceptual equivalence; therefore, the adaptation process must consider the relevance and appropriateness of the instrument's content to the target population (Haynes et al., 2026). In line with this, evaluation by experts provides a basis for assessing whether the item content remains relevant and adequate before the instrument is evaluated further through psychometric testing (Terwee et al., 2018). Thus, the content validity results in this study support that the adaptation process of the MEIM-12 went beyond literal translation and considered the conceptual appropriateness of the instrument within the context of Minangkabau society.

In addition to content appropriateness, the results of the cognitive interviews provided additional evidence regarding the respondents' response process when answering the Indonesian version of the MEIM-12. Evidence based on the response process is important because an item deemed relevant by experts is not necessarily understood and processed by respondents in accordance with the intended construct to be measured. Cognitive interviewing allows researchers to examine how respondents understand an item, access relevant experiences or information, form

judgments, and determine the response (Pereira et al., 2023). Such an examination becomes increasingly important in cross-cultural adaptation because linguistic equivalence and content relevance do not automatically guarantee that respondents from a different cultural context interpret the item as intended by the original instrument (Stefana et al., 2025). Therefore, the use of cognitive interviews in this study complements expert assessment by providing empirical evidence regarding how the MEIM-12 items are understood and used by respondents when forming their answers.

Overall, the evidence based on content validity and the response process showed that the Indonesian version of the MEIM-12 underwent an adaptation stage that considered both the conceptual appropriateness of the items and respondents' comprehension of those items. The combination of expert assessment and cognitive interviews provides complementary sources of evidence for evaluating the appropriateness of the instrument at the early stage of adaptation. However, within the framework of modern validity, evidence based on content and the response process is not by itself sufficient to support all score interpretations because evidence concerning internal structure is needed to evaluate whether the empirical relationships among items correspond to the expected representation of the construct (Rios & Wells, 2014; Sireci & Padilla, 2014). Therefore, an evaluation of the internal structure is needed to test how the MEIM-12 items empirically represent the dimensions of ethnic identity in Minangkabau society. The next stage of this study focused on testing the factor structure and reliability of the MEIM-12 scores.

Internal Structure and Reliability of the MEIM-12

Internal structure testing showed that the unidimensional model of the MEIM-12 was less able to represent the pattern of relationships among items in the sample of Minangkabau society, whereas the two-factor representation provided a better fit. This finding is in line with more recent psychometric research that continues to find support for the Exploration and Affirmation/Belonging dimensions in the MEIM across different population contexts, including young Iranian adults and ethnically diverse community samples (Fisher et al., 2020; Habibi et al., 2021). A recent study by Neumann et al. (2025) with a large sample in Aotearoa, New Zealand, also found that the two-factor structure of the MEIM had an acceptable fit, although alternative structures could also provide an adequate representation of ethnic identity. Thus, the results of this study add evidence that the two-dimensional structure of the MEIM remains relevant, but its empirical representation may show variation depending on the characteristics of the population and cultural context. The need for several residual covariances in this study also indicates that the relationships among items were not entirely explained by the two latent factors; therefore, the two-factor structure among Minangkabau society needs to be understood as a representation supported by the data of this study rather than as a structure entirely free from context-specific characteristics.

The findings of this study reinforce the evidence that the internal structure of the MEIM-12 needs to be evaluated within the context of the population in which the instrument is used. In a validation study of the MEIM-12 on 426 students in Iran with a mean age of approximately 22 years, Habibi et al. (2021) found support for a structure that distinguishes the Exploration and Commitment dimensions, indicating that ethnic identity is not always adequately represented as a unidimensional construct. Meanwhile, a recent study with a large sample in Aotearoa, New Zealand, showed that the two-factor model of Exploration and Affirmation/Belonging had acceptable fit, along with an alternative factorial representation that also showed adequate fit (Neumann et al., 2025). Overall, these findings show that the Exploration and Affirmation/Belonging dimensions remain important components in the measurement of ethnic identity, although the empirical structure of the MEIM may vary across populations. In the context of Minangkabau society, the results of this study support the two-factor structure as the most appropriate representation among the models tested. Thus, these results underscore the importance of empirically testing the internal

structure of the MEIM when the instrument is applied to different population contexts because the factor structure cannot simply be assumed to hold identically across all populations.

The need to add residual covariances for several item pairs indicated that the relationships among the MEIM-12 items were not fully explained by the two latent factors. In CFA, correlated residuals indicate the presence of systematic relationships among items that remain after the variance explained by the latent factors is accounted for (Finch et al., 2023). The addition of residual covariances in this study was based not only on statistical considerations but also on substantive considerations regarding the closeness of content between items. This is in line with the view that similarity of meaning or specific item characteristics can be a source of shared variance that is not fully explained by the main latent construct (Bandalos, 2021). Thus, the model modification provided a more appropriate representation of the data pattern without changing the basic two-factor structure of MEIM-12. However, the presence of these residual relationships still indicates that some specific characteristics at the item level need to be considered when interpreting the internal structure of the instrument.

The factor loadings in the final model showed that all MEIM-12 items had a significant relationship with the hypothesised latent dimensions. On the Exploration dimension, the factor loadings ranged from .594 to .759, whereas on the Affirmation and Belonging dimension, they ranged from .634 to .796. This range indicates that all items had a substantial relationship with the factor they measured, because all loadings were above the .50 level commonly used as one indicator of an adequate item–construct relationship (Cheung et al., 2023). Item M8 showed the strongest relationship with the Exploration dimension, whereas M7 showed the strongest relationship with the Affirmation and Belonging dimension. Conversely, M2 and M5 had relatively lower loadings than the other items in their dimensions, but both remained above the level generally considered adequate. Overall, this pattern of factor loadings provides item-level support for the two-dimensional representation of MEIM-12 in the final model.

In addition to the factor structure at the item level, the results also showed good internal consistency for both dimensions of the MEIM-12. McDonald's omega values of .806 for Exploration and .844 for Affirmation and Belonging are at a level indicating good reliability, given that omega values around .70 are generally regarded as adequate and values of .80 indicate stronger reliability (Stefana et al., 2025). Evidence of convergent validity showed a different pattern across the dimensions. The AVE value of .512 for Affirmation and Belonging exceeded the .50 benchmark, indicating that the latent factor explained slightly more than half of the average variance of its items, whereas the AVE of .456 for Exploration was slightly below that benchmark (Cheung et al., 2023). Thus, the evidence of convergent validity was stronger for the Affirmation and Belonging dimensions, whereas for Exploration, this evidence needs to be considered together with good reliability and all factor loadings being above .50. Overall, these results support the reliability of both dimensions of the MEIM-12, with stronger evidence of convergent validity for Affirmation and Belonging than for Exploration.

Overall, the results support the use of the Indonesian version of the MEIM-12 as an instrument representing two dimensions of ethnic identity, namely Exploration and Affirmation and Belonging, among Minangkabau society. Compared with the unidimensional model, the two-factor model provided a more appropriate representation of the pattern of relationships among items after modifications based on empirical and substantive considerations. All items showed adequate factor loadings, and both dimensions had good internal reliability. However, the presence of residual covariances for several items, as well as the Exploration AVE value slightly below the conventional benchmark, indicates that the psychometric evidence obtained needs to be interpreted proportionately. Thus, the results support the use of the two-dimensional scores of the MEIM-12 in this study population, but further evaluation is still needed to ensure that the measurement structure functions equivalently across different groups. Therefore, the next analysis tested

measurement invariance across sexes as additional evidence regarding the measurement equivalence of MEIM-12.

Measurement Invariance Across Sex

Measurement invariance testing across sexes showed that the measurement structure of the Indonesian version of the MEIM-12 could be maintained progressively across the male and female groups. Invariance testing through multigroup confirmatory factor analysis was carried out by applying parameter constraints sequentially, from the factor structure to the factor loadings, intercepts, and residual variances (Algner et al., 2025). The configural model showed an adequate fit, indicating that the general pattern of relationships between items and the two latent factors had a similar configuration in both groups. When the equality of factor loadings was tested using the metric model, the change in the fit indices was relatively small. The subsequent constraints on the intercepts in the scalar model and the residuals in the strict model also did not produce a substantial decline in the model fit. Based on this pattern of changes in fit indices, MEIM-12 showed support for measurement invariance up to the strict level across sexes in this study sample. This progressive approach is important because measurement equivalence must be demonstrated before differences in scores across groups can be interpreted as substantive differences in the construct being measured (Graybill et al., 2024).

Support for configural invariance indicates that MEIM-12 has the same pattern of factor structure in the male and female groups. In other words, the two dimensions identified in this study, namely Exploration and Affirmation and Belonging, are represented by a consistent pattern of relationships between items and latent factors in both groups. Configural invariance is the basis for showing that a construct has the same dimensional structure across groups (Straub et al., 2025). Furthermore, support for metric invariance indicates that each item has an equivalent relationship with the latent factor it measures in both men and women. This equality of factor loadings indicates that the relative contribution of items in representing the ethnic identity construct does not differ substantially across the sexes. Thus, the configural and metric invariance results provide evidence that the basic structure of the MEIM-12 and the relationships between the indicators and their two dimensions function consistently in men and women in this study sample.

Support for scalar invariance indicates that, in addition to having an equivalent factor structure and factor loadings, the intercept parameters of the MEIM-12 items can also be maintained equivalently between men and women. The equality of intercepts is an important stage in invariance testing because it allows comparisons of the means of the latent construct across groups to be meaningfully made (Sterner et al., 2024; Tse et al., 2024). With scalar invariance met, differences in the means of the latent construct can be interpreted as differences in the construct being measured rather than merely as a result of systematic differences in measurement parameters across groups. The results of this study thus indicate that the MEIM-12 has an adequate measurement basis for comparing the level of latent ethnic identity between men and women in the study sample. These findings extend the evidence of measurement equivalence from the similarity of structure and item–factor relationships toward the comparability of the means of the latent construct across groups, which is one of the important aims of measurement invariance testing (Protzko, 2025).

Support for strict invariance indicates that, in addition to equivalent factor structure, factor loadings, and intercepts, the residual variances of the MEIM-12 items can also be maintained equivalently between the male and female groups. In measurement invariance testing, strict or residual invariance tests the equality of the unique variances of items across groups and provides additional evidence regarding the similarity of measurement precision (Lasker, 2024; Sterner et al., 2024). In this study, the change in fit indices from the scalar model to the strict model remained small, with ΔCFI = −.001 and ΔRMSEA = −.003, as well as a non-significant scaled chi-square

comparison. This pattern indicates that constraining the residual variances to be equal did not produce a substantial decline in the model fit. Thus, the Indonesian version of the MEIM-12 showed measurement equivalence up to the strict invariance level across sexes in this study sample. This finding reinforces the evidence that the two groups not only have equivalent measurement structures and parameters but also show relatively similar levels of residual variance in the measurement of ethnic identity.

Overall, the results of the measurement invariance testing provide support for the measurement equivalence of the Indonesian version of the MEIM-12 across sexes up to the strict invariance level in this study sample. Configural invariance indicates that the two-factor structure, namely Exploration and Affirmation and Belonging, can be maintained in both male and female groups. Furthermore, the fulfilment of metric and scalar invariance indicates the equivalence of the relationships between items and the latent construct, as well as the intercept parameters across groups. Support for strict invariance provides additional evidence that the residual variances of the items are relatively equivalent in both groups. Thus, the results support the use of the MEIM-12 as an instrument with equivalent measurement parameters across sexes in the study population. However, this conclusion is limited to the sample of the Minangkabau society studied; therefore, the measurement equivalence of the MEIM-12 in other populations, age groups, or cultural contexts still needs to be evaluated through further research.

Implications

The findings also have implications for culturally responsive counselling practices. Ethnic identity is conceptualised as a multidimensional and dynamic construct involving processes of exploration and commitment and has been recognised as relevant to counselling psychology (Phinney & Ong, 2007). In this context, Indonesian MEIM-12 may provide complementary information for counsellors seeking to understand how clients explore, experience, and relate to their ethnic identity (Yap et al., 2014). Such information may help counsellors explore issues related to cultural belonging, identity negotiation, and experiences associated with family, community, and migration while considering the client's broader social and cultural context. However, MEIM-12 scores should not be used to determine the authenticity of the Minangkabau identity or to diagnose identity problems, but rather as a source of information to complement the counselling process and the client's own narrative.

For clients, the availability of an instrument that is culturally and linguistically adapted to the Minangkabau context may facilitate reflection on their experiences of ethnic identity and belonging. At the community level, the instrument may support research and community-based initiatives aimed at understanding how the Minangkabau people maintain, negotiate, and interpret their ethnic identity across different social contexts. Nevertheless, these potential applications should be understood as preliminary, as the present study evaluated psychometric evidence rather than counselling outcomes. Further research is therefore needed to examine the usefulness of MEIM-12 in diverse Minangkabau populations and in actual counselling settings.

Limitations and Recommendations

This study has several limitations. Participants were recruited through convenience and snowball sampling with an online questionnaire. The sample was relatively young (M = 29.00 years) and mostly female (58.85%), and Minangkabau people with limited internet access may be underrepresented, so the findings may not generalise to the wider Minangkabau population. The final model also needed four residual covariances, one of them between items on different factors (M3–M4), and its RMSEA stayed slightly above .08. In addition, the AVE for the Exploration dimension (.456) fell below the usual .50 criterion. Finally, measurement invariance was tested only across sex.

Future studies could address these limitations in four ways. First, more diverse samples, for example from combined online and offline data collection and with a wider spread of ages, would make the findings easier to generalise. Second, the two-factor structure should be replicated in an independent sample to check whether the residual covariances found here are specific to this sample. Third, invariance testing could be extended to other groups that matter in the Minangkabau context, such as age groups and people with and without *merantau* experience. Fourth, later work could examine validity evidence based on relations with other variables and explore how the instrument can be used in multicultural counselling practice.

CONCLUSION

This study provides preliminary psychometric evidence supporting the use of the Indonesian version of the MEIM-12 in adults in Minangkabau. First, the adaptation process showed that the Indonesian version had adequate content appropriateness and could be understood by the respondents in accordance with the context of the study population. Second, the internal structure testing supported a two-dimensional representation consisting of Exploration and Affirmation and Belonging over the unidimensional model, with adequate factor loadings and good internal reliability on both dimensions, although evidence of convergent validity on the exploration dimension was still relatively limited. Third, the MEIM-12 showed measurement invariance up to a strict level across sexes, supporting the equivalence of measurement parameters between men and women in the study sample. The instrument may also provide complementary information to support culturally responsive counselling by helping counsellors explore clients' experiences of ethnic identity and belonging within their broader social and cultural contexts. Its use in counselling should remain non-diagnostic and should be considered alongside the client's own experiences and narratives.

AUTHOR CONTRIBUTIONS STATEMENT

IS: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Writing-Original Draft, and Visualization. CY and SL: Conceptualization, Methodology, Supervision, Validation, and Writing-Review and Editing. HS & IO: Data Analysis, Manuscript Development. All authors have read and approved the final manuscript.

REFERENCES

Algner, M., Fay, D., & Lorenz, T. (2025). Validation and measurement invariance of the german and spanish gender bias scale for women leaders. *Journal of Business and Psychology*, *40*(1), 155–178. https://doi.org/10.1007/s10869-024-09946-7

Ariely, G. (2019). The nexus between globalization and ethnic identity: A view from below. *Ethnicities*, *19*(5), 763–783. https://doi.org/10.1177/1468796819834951

Badan Pusat Statistik. (2024). *Profil suku dan keragaman bahasa daerah hasil long form sensus penduduk 2020*. Badan Pusat Statistik.

Bandalos, D. L. (2021). Item meaning and order as causes of correlated residuals in confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(6), 903–913. https://doi.org/10.1080/10705511.2021.1916395

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2023). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, *41*(2), 745–783. https://doi.org/10.1007/s10490-023-09871-y

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16

Finch, W. H., French, B. F., & Hazelwood, A. (2023). A comparison of confirmatory factor analysis and network models for measurement invariance assessment when indicator residuals are correlated. *Applied Psychological Measurement*, *47*(2), 106–122. https://doi.org/10.1177/01466216231151700

Fisher, S., Zapolski, T. B., Wheeler, L., Arora, P. G., & Barnes-Najor, J. (2020). Multigroup ethnic identity measurement invariance across adolescence and diverse ethnic groups. *Journal of Adolescence*, *83*(1), 42–51. https://doi.org/10.1016/j.adolescence.2020.07.006

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312

Graybill, E., Barger, B., Salmon, A., & Lewis, S. (2024). Gender and race measurement invariance of the strengths and difficulties questionnaire in a U.S. base sample. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1310449

Habibi, M., Etesami, M. S., Taghizadeh, M. A., Akrami, F. S., & Garcia, D. (2021). Psychometric properties of the multi-group ethnic identity measure (MEIM) in a sample of iranian young adults. *PeerJ*, *9*, e10752. https://doi.org/10.7717/peerj.10752

Handrianto, C. (2017). The roles of matrilineal system towards integrating religious and cultural values in minangkabau community. *Jurnal Ilmiah Peuradeun*, *5*(3), 373–386. https://doi.org/10.26811/peuradeun.v5i3.170

Haynes, S. N., Matsui, M., & Kaholokula, J. K. (2026). Content validation and cultural adaptation are necessary to interpret evidence from psychological assessment instruments imported across cultures. *Current Opinion in Psychology*, *70*, 102307. https://doi.org/10.1016/j.copsyc.2026.102307

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

International Test Commission. (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). International Test Commission. https://www.intestcom.org/

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2026). *semTools: Useful tools for structural equation modeling*. https://cran.r-project.org/package=semTools

Kato, T. (1982). *Matriliny and migration: Evolving minangkabau traditions in Indonesia*. Cornell University Press.

Lasker, J. (2024). Measurement invariance testing works. *Applied Psychological Measurement*, *48*(6), 257–275. https://doi.org/10.1177/01466216241261708

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Neumann, D., Yao, E. S., Feng, L., Taufa, S., Liang, R., Kingi, T. K., Atatoa Carr, P., Fenaughty, J., & Paine, S.-J. (2025). Exploring ethnic identity in early adolescence: Psychometric evaluation of the multigroup ethnic identity measure (MEIM) in an ethnically diverse population-based sample of young people in aotearoa new zealand. *Self and Identity*, *24*(6), 688–713. https://doi.org/10.1080/15298868.2025.2517189

Nugraha, Y. C. (2022). Indonesian international students' identity post mobility: How mobility affects international students beyond education. *Populasi*, *30*(2), 1–14. https://doi.org/10.22146/jp.80181

Parker, L. (2011). Theorizing adolescent identities: Gender dynamics and differences in west sumatra. *The Asia Pacific Journal of Anthropology*, *12*(1), 6–26. https://doi.org/10.1080/14442213.2010.536708

Pereira, G. U., Beal, G. M., & Oliveira, M. Z. de. (2023). The cognitive interview in the adaptation and validation of a psychometric instrument. *Avaliação Psicológica*, *22*(1), 2–9. https://doi.org/10.15689/ap.2023.2201.18944.01

Phinney, J. S. (1992). The multigroup ethnic identity measure: A new scale for use with diverse groups. *Journal of Adolescent Research*, *7*(2), 156–176. https://doi.org/10.1177/074355489272003

Phinney, J. S., & Ong, A. D. (2007). Conceptualization and measurement of ethnic identity: Current status and future directions. *Journal of Counseling Psychology*, *54*(3), 271–281. https://doi.org/10.1037/0022-0167.54.3.271

Posit Team. (2026). *RStudio: Integrated development environment for R*. Posit Software, PBC. https://posit.co/

Protzko, J. (2025). Invariance: What does measurement invariance allow us to claim? *Educational and Psychological Measurement*, *85*(3), 458–482. https://doi.org/10.1177/00131644241282982

Revelle, W. (2026). *psych: Procedures for psychological, psychometric, and personality research*. https://cran.r-project.org/package=psych

Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*, *26*(1), 108–116. https://doi.org/10.7334/psicothema2013.260

Rosseel, Y., Jorgensen, T. D., & De Wilde, L. (2025). *lavaan: Latent variable analysis*. https://doi.org/10.32614/CRAN.package.lavaan

Sireci, S., & Padilla, J.-L. (2014). Validating assessments: Introduction to the special section. *Psicothema*, *26*(1), 97–99. https://doi.org/10.7334/psicothema2013.255

Stefana, A., Damiani, S., Granziol, U., Provenzani, U., Solmi, M., Youngstrom, E. A., & Fusar-Poli, P. (2025). Psychological, psychiatric, and behavioral sciences measurement scales: Best practice guidelines for their development and validation. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1494261

Sterner, P., Pargent, F., Deffner, D., & Goretzko, D. (2024). A causal framework for the comparability of latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *31*(5), 747–758. https://doi.org/10.1080/10705511.2024.2339396

Straub, N., Goretzko, D., & Sterner, P. (2025). Assessing measurement invariance with exploratory factor analysis trees: A practical guide. *European Journal of Psychological Assessment*, 1015-5759/a000898. https://doi.org/10.1027/1015-5759/a000898

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A delphi study. *Quality of Life Research*, *27*(5), 1159–1170. https://doi.org/10.1007/s11136-018-1829-0

Tse, W. W.-Y., Lai, M. H. C., & Zhang, Y. (2024). Does strict invariance matter? Valid group mean comparisons with ordered-categorical items. *Behavior Research Methods*, *56*(4), 3117–3139. https://doi.org/10.3758/s13428-023-02247-6

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2026). *dplyr: A grammar of data manipulation*. https://doi.org/10.32614/CRAN.package.dplyr

Yap, S. C. Y., Donnellan, M. B., Schwartz, S. J., Kim, S. Y., Castillo, L. G., Zamboanga, B. L., Weisskirch, R. S., Lee, R. M., Park, I. J. K., Whitbourne, S. K., & Vazsonyi, A. T. (2014). Investigating the

structure and measurement invariance of the multigroup ethnic identity measure in a multiethnic sample of college students. *Journal of Counseling Psychology*, *61*(3), 437–446. https://doi.org/10.1037/a0036253